ELEMENTARY ECONOMICS

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN AND CO., Limited
LONDON · BOMBAY · CALCUTTA · MADRAS
MELBOURNE

THE MACMILLAN COMPANY
OF CANADA, Limited
TORONTO

ELEMENTARY ECONOMICS

by

Fred Rogers Fairchild
KNOX PROFESSOR OF ECONOMICS

Edgar Stevenson Furniss
PELATIAH PERIT PROFESSOR OF SOCIAL SCIENCE

Norman Sydney Buck
PROFESSOR OF ECONOMICS

YALE UNIVERSITY

VOLUME II
Fourth Edition

THE MACMILLAN COMPANY
NEW YORK

FOURTH EDITION COPYRIGHTED, 1939,
BY THE MACMILLAN COMPANY.

All rights reserved — no part of this book may be reproduced in any form without permission in writing from the publisher, except by a reviewer who wishes to quote brief passages in connection with a review written for inclusion in magazine or newspaper.

Printed in the United States of America
Set up and electrotyped. Published May, 1939.
Reprinted November, 1939, June, 1940.
May, 1944 September, 1945. June, 1946

Copyright, 1925 and 1926,
By Fred Rogers Fairchild
First, second, and third editions copyrighted and published 1926, 1930, and 1936
By The Macmillan Company

CONTENTS

PART V. TAXATION AND THE BUSINESS OF GOVERNMENT

PART VIII. PROBLEMS OF LABOR

PART IX. PROPOSED CHANGES OF THE ECONOMIC SYSTEM

PART V

TAXATION AND THE BUSINESS OF GOVERNMENT

XXIX

PRINCIPLES OF GOVERNMENT INCOME

Government and the economic organization. No account of the economic organization would be complete without attention to the part played by government[1] in the production, distribution, and consumption of wealth. The existence for government is familiar to everyone, and virtually everyone recognizes its necessity. In fact we incline so much to take government for granted that it may not be amiss to lay the foundation for the study of this part of our subject by noting very briefly some of the more important economic aspects of government.

In modern times the greatest consumer of goods and services is the government. In 1938 the economic activity of the people of the United States brought forth a total income which is estimated as worth 62¼ billion dollars. The government (including all grades, federal, state, and local) collected and expended about 17 billions, about 27 per cent of the total national income. In taxes alone there were collected about 14 billions. In the prosperous year 1929 when the national income was estimated at 79½ billion dollars, government collected and expended 11⅝ billions.[2]

On the side of production, the government is ever present. For example the United States Post Office furnished services in the fiscal year ending June 30, 1938, which cost over seven hundred and seventy-six million dollars. In this same year the Tennessee Valley Authority cost 42 millions, and the Panama Canal, 11 millions; on public works the federal government spent the huge sum of 1,849 millions. But the chief contribution of government to production is

[1] The term "government" is used in this book to include all grades of government of any particular community, for example in the United States of America the federal government, the state governments, and the governments of counties, towns or townships, cities, school districts, and all other political jurisdictions. It is rather common in popular speaking and writing to refer to the United States federal government as "the government" in distinction from the states and the local jurisdictions. The term is not so used in this book.

[2] United States Census, *Financial Statistics of State and Local Governments, 1932*, 1935; National Industrial Conference Board, *Economic Record*, Aug. 31, 1939.

indirect, consisting in such services as protection from outside foes, enforcement of law and order among the people, promotion of the public health, public education, etc.

In the field of distribution, the government is no less omnipresent and important. We need only to recall that in the year 1936 over one-fourth of the total income of the American people was collected and redistributed by government. In the year 1936 the various governments in the United States (federal, state, and local) paid nearly 6 billion dollars in compensation to 3,625,000 employees, not including work relief employees.[1] These employees comprised almost 9 per cent of all persons employed in the country (not including unpaid family labor).

The functions of government. Lest they be overlooked in the discussion of the more specialized and controversial governmental activities, a brief summary of the normal functions of government may not be out of order. The primary functions of government include (1) those of a defensive nature: defense against outside foes, requiring army, navy, etc.; justice and security, afforded by police service, courts, and penal institutions; regulation of private industry, in particular monopolistic industries like the railroads and those which are so complicated as to be only imperfectly understood by the public, such as the banks and insurance companies; protection against disease and accident; maintenance of moral standards; and protection against fire, storms at sea, and certain other forces of nature; and (2) developmental functions: education; religion; facilities for private industry, especially highways and bridges, harbors, the monetary system, information services, etc.; and development and conservation of the nation's natural resources. In addition to the foregoing, which are peculiarly governmental functions in modern times, government may enter the industrial field, where private enterprise generally holds sway, and we have a considerable variety of government industries.

Government income. To meet the heavy expenditures necessitated by the performance of its functions, the modern government

[1] United States Bureau of Foreign and Domestic Commerce, *National Income, 1929–1936*, 1937. For further light on this topic, cf. William E. Mosher and Sophie Polah: "Public Employment in the United States," *National Municipal Review*, January, 1932, pp. 51–75.

must have an equally large income. In modern times there are just four important sources from which government income is to be obtained and which now demand our study. These are (1) industrial earnings (prices), (2) taxes, (3) fees and assessments, and (4) loans.[1]

The finances of government industry. When a government engages in a business enterprise, such as the United States Post Office or the water supply system of a city, it sells its product (goods or service) at a definite price to the consuming public much as any private business concern does. The citizen may take the offered service or not as he wishes, and if he does not take it he makes no payment. Payment is voluntary and in proportion to the amount of the service received. For example the government offers, through the post office, to transmit letters and parcels If the citizen chooses to take advantage of this service he pays the stated price; otherwise he pays nothing to the post office. In these respects government industry is similar to private business

In contrast with the private entrepreneur, the government's purpose in its industrial undertakings is ordinarily, not the making of a profit, but the furnishing to the people of a service which it is believed can be better provided by the government than by private enterprise. Thus the post office exists, not to make profits for the government, but to afford the people facilities for the widespread communication of messages and information and the cheap and easy transmission of merchandise in a manner which it is believed could not be obtained from private agencies.

In general, governments seek to make their industrial enterprises pay their way; prices are fixed so that the receipts shall just about equal the expenditures in each particular industry. There are, it is true, many exceptions. Sometimes a government reserves for itself a monopoly of a certain business and runs the industry frankly for the sake of making a profit for the public treasury. The French tobacco monopoly is an example This result is scarcely possible unless the government enjoys a monopoly position; other-

[1] For the sake of completeness and to avoid possible misunderstanding it may be well to mention certain minor sources of government income whose investigation the limits of this book will preclude The most important of these are gifts, forfeiture, reversion, escheat indemnity confiscation, fines and penalties

wise the competition of private business firms will prevent the government industry from securing a profit. On the other hand, some government industries do not secure revenues sufficient to meet their expenses, thus showing a deficit, which has to be made up out of the general treasury. The United States Post Office is regularly in this state. In any event each government industry yields industrial earnings which are earmarked to meet the expenses of that particular industry.

Financing the primary functions of government: General character of revenue. If now we pass from its industrial undertakings to the primary functions of government, we shall find certain differences in their financing. The primary functions, such as defense, justice and security, protection against disease, education, etc., are undertaken by the government in order to benefit the people as a whole, without the purpose of giving particular benefit to particular individuals. For example the state and local governments of the United States provide public schools and universities, at great expense, not primarily in order to benefit those persons who want an education for themselves or their children, but because it is believed that a prosperous and enduring democracy is not possible unless there is widespread education among the people. Education is provided therefore to benefit the whole people, not for the sake of those particular individuals who attend the public schools.

In all these cases therefore no attempt is made to pay the cost of the service by charging a price and so collecting from each citizen in proportion to the benefit he has received. The government does not sell education as it sells postal service; education is offered freely to all. A may have one child in the public school, while B has half a dozen. No account is taken of these facts in determining the respective amounts which A and B shall pay toward the support of the schools. Furthermore no citizen is able to avoid support of the public school by declining to avail himself of its services. C may send all his children to private schools and there pay for their education, and D may have no children at all. Neither will find it possible to cite these facts as reasons why he should not contribute to the cost of the public schools. The burden of furnishing the money necessary to pay the cost of public education, as of the other

primary functions of government, is distributed among the people according to certain rules, which we shall presently study, but there is no necessary relation between the amount of any person's contribution and the service received by him, nor may any person avoid contributing by declining to avail himself of the government's service. Contribution toward the cost of the primary functions of government is compulsory.

It follows that no attempt is made in the case of the primary functions of government to obtain in connection with each function a revenue earmarked to meet the expenditures of that particular function. As a general rule, little or no revenue comes in to the government in connection with the performance of any particular function, and the bulk of the revenue is received without any reference to the particular functions of government for which it is to provide.

Taxes. The foregoing analysis of the purpose of the government's primary functions and the financial contrast between such functions and the government's industrial undertakings has introduced the principal characteristics of taxation; namely, (1) compulsion, (2) devotion to the general welfare, and (3) apportionment without reference to individual benefit. It will be found helpful to have a precise definition of a tax, for which the following will satisfactorily serve our purpose: *A tax is a compulsory payment by a person to the government, destined to defray the cost of government services performed for the common benefit, and paid without reference to individual benefit.* It is upon taxes that modern governments principally rely for obtaining the money necessary to defray the cost of performing their primary functions, and taxes make up the bulk of the income of most modern governments.

Fees. The primary functions of government are performed, as has been shown, for the common benefit. Nevertheless it often happens that, in performing its primary functions, the government incidentally and unavoidably gives special benefits to particular individuals. Consider for example the granting of patents for inventions. The purpose of this service of government is not to give a special benefit to the inventor; if such were its purpose it would be a case of favoritism, repugnant to the spirit of democratic gov-

ernment. The real purpose of the patent is to promote the public welfare by encouraging the discovery and disclosure of useful new devices and by offering inducement to capitalists and entrepreneurs to develop and market them It is for this reason that the inventor is given for a time a privileged position of monopoly with respect to his discovery. Nevertheless this is a material benefit conferred upon the inventor, and for that reason the government imposes upon him a special charge, justified by and — somewhat roughly it must be confessed — related to the particular benefit which he is presumed to have received. Such a payment is called a fee.

Fees arise also from somewhat different circumstances. There are certain occupations which cannot be left to the free and unrestricted operation of all persons without danger to the public interests. Such are the businesses of banking, insurance, taxicab operation, peddling, etc. The public oversight of such private enterprises necessitates government expenditure, and the person who chooses to enter such an occupation is thereby causing expense to the government. The government therefore charges such person a fee, covering part or all of the cost imposed upon it by his engaging in the regulated occupation. Fees are thus imposed upon those who engage in any activity over which some government regulation has been found necessary One cannot drive an automobile, or own a dog, or enter the marriage state, without a government license, and in each case he is charged a fee to cover the cost which his act has imposed upon the government.

We now have in mind the nature of the fee, which may be defined in precise terms as follows: *A fee is a payment by a person to the government on account either of a special benefit received from the government or a special cost imposed upon the government in connection with a government service performed for the common benefit.*

Special assessments. Suppose that the city authorities decide to pave a certain street There will generally result a material advantage to the owners of land fronting that street on account of the rise in the value of such land. The same rise in land values follows the introduction of a sewer system, street curbs, sidewalks, and other improvements. The city government undertakes a new street pavement, not from any desire to benefit particularly the owners of the

land, but because the interests of all the people of the city will be promoted by a better surface on that street. Here then is a perfect occasion for a fee and, in America at least, a fee is usually charged, which is none the less a fee in spite of its common designation as a *special assessment*. This is the definition: *A special assessment is a particular kind of fee, imposed when the special benefit consists in an enhancement of the value of land.* The amount of the special assessment may be anything up to the total addition to the value of the land caused by the government improvement. The special assessment furnishes a most interesting example of the fee, since there is here usually an attempt to relate the amount of the payment to the value of the special benefit conferred much more closely than in the case of fees in general.

Names may be misleading. In distinguishing prices, fees, and taxes, we must have regard to their real nature and not be misled by mere names. For example a fee is never greater than the value of the special benefit received or the special government cost occasioned by the fee-payer. If this limit is exceeded the payment is, at least to the extent of the excess, a tax. This is usually the case with the automobile operator's "license fee" collected by the states of the United States. On the other hand, consider the tax upon gasoline, which has come into recent prominence in the United States. This tax is imposed by the state at so many cents per gallon, and the proceeds are devoted chiefly to the maintenance and repair of the highways. When the facts are thus, this so-called "tax" is really a fee. It is a payment which may be justified either on account of the benefit received from the public highways or the cost imposed by wear and tear upon the highways, and it apportions with reasonable accuracy the size of the payment to the amount of such benefit or cost.

Distribution of the tax burden: Equal distribution. Almost the entire revenue required to meet the cost of government's primary functions comes in the form of taxes, which are exacted from the people with little or no regard to particular benefits received. Great importance therefore attaches to the distribution of the tax burden; that is, the determination of how much each person must pay for support of the general services of government.

It may perhaps serve to clarify the first steps of the investigation of this problem if we stop for a moment to consider the possibility of equal distribution. This will be recognized as the rule which usually determines the contributions of the members for support of the general services of an ordinary club. Let us see how it would work in the United States. The total tax bill of the people of the United States in the year 1936 was about ten and a half billion dollars. With a population of about 128 millions, the share of each person — man, woman, and child — would be about eighty-two dollars. At this rate a family of five (father, mother, and three children) would have an annual tax bill of $410. If the man's wages or salary and other income were $1,500, his tax bill would be more than a quarter of his income. Another family of five with an income of $10,000, in paying the same amount, would pay only about one twenty-fifth of their income, while the very wealthy would find their tax burdens quite insignificant. To the very poor the burden would be absolutely crushing, if indeed they were able to pay at all.

Now quite apart from any question of justice, it is evident that this rule would be unworkable, since there are large numbers of the people who could not possibly pay at the rate of eighty dollars a year for each taxpayer and all those dependent upon him. And whatever one's standard of justice may be, it is safe to say that virtually everyone would condemn this plan as flagrantly unjust. The important conclusion is that public opinion tolerates — even demands — a plan of tax distribution which makes some persons pay more than others. Having cut loose from the base of equal distribution, what shall be our rule of apportionment?

Benefit or cost. There have been those who believed that the costs of government should and could be distributed on the basis of the benefits received by the individual members of the group, or on the somewhat similar basis of the cost to the government of the service rendered to each person. These are in general the rules that determine prices charged in business dealings, and we have seen that they apply also to industrial enterprises conducted by government. But when we come to the general services of government, performed for the common good rather than for the particular bene-

fit of individuals, there is generally no way of measuring either the benefit or the cost to each person. How should we apportion to each citizen of the United States the benefit received by him from the activities of the army and the navy? What has the Supreme Court of the United States done for him? What part of the President's salary should he pay on the basis of benefit received? What part of the cost of the United States navy is imposed upon it by a peaceful and law-abiding Illinois coal miner who has never been outside his home state? And if it cost the navy and marines half a million dollars to save a citizen who fell into the clutches of a tribe of bandits, shall he be made to pay the cost? And if he be a poor man, shall he be protected only to the extent of his ability to reimburse the government for the cost? Even were there some ascertainable relation, the government could collect only from those who were able to pay, though its general services must be rendered for all without discrimination.

It should be clear that neither benefit nor cost can serve as a general rule for the apportionment of taxes. At the same time these principles need not be discarded entirely. Just as they serve well as the bases for fees and assessments, so they have a certain bearing at least upon the obligation to pay taxes, if not upon the amount of the individual contribution. This will come up again in our study of particular taxes.

Ability to pay: Meaning. The almost instinctive rejection of the plan of equal distribution of the tax burden is based upon the common feeling that those who have much should pay more than those who have little, that there should be at least some approach to equality of sacrifice, that people should pay taxes in proportion to their ability to pay. There is no doubt a lack of precision in these notions, but they are widely held and with sufficient clearness to indicate another rule for tax apportionment; *i.e.*, the principle of ability to pay, which may be paraphrased accurately enough as the principle of equal sacrifice.

It is not our purpose to undertake a philosophical analysis of the principle of payment according to ability; its broad meaning will be clear enough for our present purpose. As so understood, the ability principle has on the whole commended itself to modern students of

taxation and to modern public opinion as the basic principle of tax apportionment, to be supplemented by the principles of benefit and cost so far as they may be applicable and to be modified on occasion by other principles of a non-fiscal character. It must never be forgotten however that the complex practical problems of taxation are not to be solved by the mechanical application of any set of general principles. Modifications of general rules and resort to other principles will be found necessary to meet the exigencies of particular kinds of taxes.

How is the ability to pay measured? The collection of taxes being a practical financial operation, there is necessity of a financial yardstick with which to measure taxpaying ability. For all practical purposes, just three measures are available: (1) the possession of wealth (or property), (2) the receipt of income, (3) consumption, or expenditure. We may fairly assume that to the student who has reached this point in the study of economics the appropriateness of these measures will be evident.

In the application of the property and income yardsticks to the measurement of taxable ability, it is generally the net amount, rather than the gross, that more accurately measures taxpaying ability. A and B may each own wealth of a value of $50,000. If A has debts of $30,000, representing property rights of others against him and his wealth, while B has no such claims against him, it is evident that the taxpaying abilities of the two, so far as indicated by wealth or property, are not the same, and that A's ability is to be measured by his net possessions of $20,000. Speaking strictly and as we have previously defined the terms, it is not wealth, but net property, that is the real measure. Net property is the difference between all the taxpayer's possessions (including property rights against the wealth and persons of others) and the property rights of others against him and his wealth. Likewise in using the income measure, it is net income; *i.e.*, the difference between gross income and the expenses necessary to secure such income, that is the real measure of taxable ability.

Proportional taxation. There arises at once the question whether the relation between wealth, income, or consumption and taxable ability is one of simple proportion or is something different. Sup-

pose that A has a net income valued at $2,000, while B's is worth $4,000, and C's, $40,000. Does this indicate that B should pay twice as much as A, and that C should pay ten times as much as B and twenty times as much as A? If so we have proportional taxation, to be accomplished by a uniform rate. Suppose the rate is 4 per cent. Then we have this result, in which the tax payments are in direct proportion to the respective net incomes:

Person	*Net income*	*Tax rate*	*Amount of tax*
A	$ 2,000	4%	$ 80
B	4,000	4%	160
C	40,000	4%	1,600

Progressive taxation. If instead of making the tax payments directly proportional to the base, they are made to increase at a rate faster than the increase in the base, we have progressive taxation. For example suppose a scheme of income tax rates such as this:

On the first $1,000 of net income			1%
On the part of the net income:			
over $	1,000, up to and including $	2,000	2%
over	2,000, up to and including	3,000	3%
over	3,000, up to and including	5,000	4%
over	5,000, up to and including	10,000	5%
over	10,000, up to and including	20,000	6%
over	20,000, up to and including	40,000	7%
over	40,000, up to and including	60,000	8%
over	60,000, up to and including	80,000	9%
over	80,000, up to and including	100,000	10%
over	100,000		11%

Now A, with a net income of $2,000, will pay a tax of 1 per cent of his first $1,000 plus 2 per cent of his second $1,000; *i.e.*, $10 + $20 = $30. Similarly B's tax would be calculated thus:

1% of the first	$1,000,	$ 10
2% of the second	1,000,	20
3% of the third	1,000,	30
4% of the fourth	1,000,	40
Total tax on	$4,000,	$100

C's tax would amount to $2,390. These results may be tabulated thus:

Person	*Net income*	*Tax rate*	*Amount of tax*	*Ratio of tax to net income*
A	$ 2,000	1%-2%	$ 30	1.5%
B	4,000	1%-4%	100	2.5%
C	40,000	1%-7%	2,390	5.975%

Under such a set of rates, B's tax would be more than three times that of A, though his income is only double A's. C's income is twenty times A's, but his tax is practically eighty times as great. This is the result of progressive taxation and is accomplished, be it noted, by giving up the uniform rate in favor of a rate which increases with increases in the base.

It is probably not possible to prove scientifically that progressive taxation is required by the principle of ability. For that we should need a calculus for the measurement of sacrifice which thus far does not seem to have been discovered. However there is little doubt that the consensus of modern opinion, expert and general, inclines to the support of progression.

If there is no scientific proof of the correctness of progressive taxation, still less can there be any scientific determination of the exact scale of progressive rates required to apply the ability principle. In actual legislation the rates of progressive taxes are determined principally by rule of thumb, guided by various principles and interests, harmonious or conflicting. This is a matter to which we shall later return.

Regressive taxation. It is possible to apply a principle opposite to progression, by fixing rates which decrease as the tax base increases. This is called regressive taxation. There are few if any who defend this rule as an accurate application of the ability principle, and it never appears consciously embodied in tax legislation.

The problem of taxing according to ability. There is no one perfect yardstick for the measurement of taxpaying ability. A general property tax which reached the entire net property rights of every natural person would be an accurate measure so far as it went. It would be defective in that it would fail to take account of the ability obviously conferred by the receipt of income from other sources than property, such as salaries, professional fees, and wages. Suppose that A and B each owns property (corporate securities, let us say) yielding an annual income of $5,000 and worth $125,000 and that, while A has no other income, B is a lawyer earning $10,000 in fees. Obviously a property tax alone, exacting equal contributions from A and B, would miss widely the mark of taxation according to ability.

On the other hand, the commonly held notion that income is the one perfect measure of taxable ability can easily be shown to be erroneous. Suppose C and D each to be possessed of $1,000,000; C invests in corporate securities yielding an annual income of $40,000; D invests three-fourths of his fortune in idle real estate and derives only an income of $10,000 from the remaining $250,000 invested in corporate securities. The income tax alone will obviously fail to reach the taxpaying ability of D. And it is no sufficient answer to point out that if D's real estate venture proves successful, he will eventually have an increased income, upon which he will be taxed. The venture may prove unsuccessful. And in any case D is not today contributing to the present cost of government according to his ability. Each of these men was free to invest his fortune as he chose. Assuming neither to have any other source of income, their taxpaying abilities before they made their respective arrangements were equal. In this case the property tax would come closer to apportioning taxation according to ability to pay.

Consumption taxes likewise furnish only an imperfect measure of taxpaying ability. A general consumption or sales tax obviously takes more from the rich than from the poor and so recognizes the principle of ability. But such a tax alone is a very imperfect device for reaching taxpaying ability. It fails to reach the ability represented by property that is not yielding a present income (as in the case of D), and even when there is income, it fails to take account of that part of the income that is saved. Suppose E and F are neighbors with about the same needs and the same scale of expenditures, each spending around $3,000 a year, although E's income is $6,000 while F's is $3,000. E's taxable ability is double that of F — if not more — and is clearly not reached to its full extent by a tax on consumption or spending, which takes only the same amount from him as from F.[1] The relation between tax payment and ability is still more remote when the consumption tax is, as is usual, levied upon a selected list of commodities or services instead of upon all expenditures.

[1] It may be noted that, if income be defined strictly as the actual benefits or services rendered by wealth or free persons (cf. Chapter I), savings are not a part of net income, and the income tax also would fail to reach the full taxpaying ability of E.

The three measures of taxpaying ability differ as to their adaptability to progression. Consumption taxes are seldom imposed directly upon the consumer; they are usually levied against manufacturers and dealers, through whose hands the goods pass. It is impossible in such case to use progressive rates in relation to the total consumption of the individual consumer. Even where the tax is imposed directly upon the purchaser, as in some of the recent state sales taxes, it would be impracticable to use progressive rates. The property tax might conceivably be progressive. But there are serious difficulties, both theoretical and practical, and the property tax is virtually always a proportional tax. It is only the income tax that lends itself readily to progression.

In discussions of the problem of tax apportionment, use is often made of somewhat different concepts of progressive and regressive taxation. According to such concepts, taxation is considered progressive when the amount paid by the individual increases at a rate faster than the increase in some assumed general or composite measure of economic power; roughly "the tax bears more heavily upon the rich than upon the poor." In this sense the progressive income tax (even the proportional income tax with its ordinary exemptions, credits, etc.) is progressive. So to a degree is the property tax. In the first instance this tax generally falls lightly or not at all on the poor. However as will appear more clearly after our study of shifting and incidence in the next chapter, a part of the ultimate burden of the property tax is passed on to many who do not pay the tax directly. Consumption taxes, on the other hand, are considered regressive (even though they are imposed at uniform rates), because, in the first place, they are usually levied upon articles of wide general consumption which call for a greater share of the poor man's expenditure than of the rich man's, and secondly they impose no burden on that part of income which is saved, a feature of which the poor can take only slight advantage. This concept of progression, though lacking in precision, is useful when passing judgment upon a tax system.

Benefit, cost, and ability. General and special business taxes, measured more or less accurately by the amount of business done or the amount of business profits received, are quite common. Such

taxes are well known in Europe. In America they are found in the business license taxes of many of the states, particularly in the South, in the liquor licenses commonly imposed by towns and cities, and in most of the special taxes on corporations. When imposed in addition to, not in lieu of, the regular taxes upon either the property or the income of the business enterprise or upon the owner on account of the property employed in the business or the income derived from it, business taxes can with difficulty be related to the principle of ability, but must generally be frankly acknowledged as applications of the benefit or the cost principle. The assumption is that government gives special protection to property employed in business and special services to those persons, natural or corporate, so engaged, or that the cost of government is especially increased on account of business activities of the people.

The poll tax, being usually a uniform contribution from all individuals, has no relation whatever to ability, but is related to the benefit principle, on the somewhat dubious theory that protection of life is granted to all and is of equal value to all.

The fact that business taxes, poll taxes, and possibly certain other forms of taxation are not in full harmony with the principle of ability is not to be taken as an *a priori* condemnation of such taxes, but is rather an indication of the truth of our earlier statement that no one principle can be relied upon to shape all the taxes which any government finds it necessary to impose. It must further be remembered that we cannot pick out a single tax from a tax system and criticize it as though it stood alone. Final judgment must be upon the tax system as a whole. Thus it is possible to combine certain taxes measured according to benefit or cost with other taxes measured according to ability, as well as to combine proportional consumption taxes with progressive income taxes, inheritance taxes, etc., in such manner that the whole tax system may have reasonable conformity to the principle of ability.

Lawful escape from taxation. No one likes to pay taxes. It is only human nature that every taxpayer should use all lawful means to make his taxes as light as possible and that some taxpayers should resort to measures that are not lawful. Certain modern

taxes, such as the property tax, the income tax, and the inheritance tax, are so involved with complexities that it is often difficult for the taxpayer to know just what his legal obligation is. For example there can be no doubt that, under the United States income tax, many persons are paying more than their lawful taxes, through misunderstanding of the law or mere carelessness. Under such circumstances the use by some taxpayers at least of every lawful means, including the employment of tax experts and lawyers, in order to avoid paying more than is legally due, is inevitable. Again a person may escape a heavy tax upon his property by selling his real estate and moving to another town where the tax rate is lower, or a man may avoid an inheritance tax by giving his property to his intended heirs during his lifetime. Finally it is often possible to shift the burden of a tax to other persons through increasing the price charged for some commodity or service. Put in general terms, it may be possible to escape a tax burden that might otherwise have to be borne (1) by exercising care not to pay more than is legally due, (2) by putting one's self in such position that the tax law is not applicable, or (3) by using the power which a tax may give to raise the price of something sold.

Now so far as all of these expedients are legal, it can raise no ethical controversy to assert that they are entirely justifiable. For the first the legislature itself is principally responsible, through enacting tax laws too complicated or obscure for the understanding of the ordinary individual. In the second case, if it appears that taxpayers are finding it easy so to adjust their residences or businesses or possessions or other circumstances as to put themselves outside the application of a tax, thus defeating the purpose of the law, that again must be due to defects in the tax itself, for which the lawmaking body alone is responsible. The case of shifting the burden to others may indicate either failure of the law to place the final burden upon those whom the legislature intended to reach or simply the result foreseen and intended by the legislature. If the first, then the miscarriage is not to be ascribed to any dereliction, legal or moral, on the part of the taxpayer, but simply to lack of understanding of economic laws on the part of the lawmakers. As to the second, we have the simple and common case of indirect

taxation, which will be examined presently, but which requires no further comment in this connection.

Tax evasion. In popular discussion there often appears a certain failure to distinguish between such legal and proper escape from the burdens of taxation and the evasion of taxes legally due. However much this popular confusion may be excused by the obscurity of tax laws and the complexities of the whole modern tax situation, the student must always keep the distinction in mind. Escape from taxation may be lawfully accomplished in a variety of ways; the term "evasion" had better be reserved for violations of the tax laws.

When tax evasion occurs in sporadic instances, the situation is similar to any other case of violation of law and is of no special concern to the economist or the tax student. When, as is sometimes the case, there is wholesale evasion, the responsibility must be either upon the legislature for enacting a law incapable of enforcement or upon the administrative personnel. To ascribe it to a low state of public morality does not face the problem. The legislature has to deal with human nature as it is and is responsible for enacting laws which can be enforced and for providing the necessary machinery for efficient administration. We shall find instructive applications of these various principles when we come to the study of the different kinds of taxes.

The shifting of taxes: Meaning. Taxes are *shifted* when the person upon whom the tax has been imposed finds that the presence of the tax has given him the power, which he did not before possess, to raise the price of some commodity or service which he has for sale, thereby reimbursing himself in whole or in part for his tax payment and to that extent passing the burden on to someone else. An example familiar to everyone is the United States internal revenue tax upon the manufacturers of tobacco products. A comparatively heavy tax is thus imposed, but the manufacturers, who are the immediate taxpayers, regularly pass the burden of the tax on to the consumers of tobacco through an increase of the price. The degree of shifting which is possible varies, according to the nature of the tax, all the way from those whose burden is entirely or almost entirely shifted, such as the tobacco tax just cited, to

those whose shifting is impossible, such as a tax upon monopoly profits or upon economic rent. Shifting and *incidence* (meaning the final resting place of the tax burden) must never be overlooked when considering the problem of the distribution of the tax burden or when criticizing any particular tax. Obviously the person who finally bears the burden is more to be considered than he who pays only to shift the burden to others.

Direct and indirect taxes. Taxes are usually put into two classes according as the burden is or is not normally shifted by the one from whom the government collects the tax. Direct taxes are those whose burden remains on the whole upon those from whom the tax is originally collected. Indirect taxes are those whose burden is normally shifted. Since the possibility of shifting may be present only partially and in various degrees, it is not always possible to draw a sharp exclusive line between the two classes of taxes. However the characteristics of the more important kinds of taxes are so clearly marked that there is usually little difficulty in placing any given tax in its proper class. This distinction, be it noted, is an economic one. It does not agree precisely with the legal definition of a direct tax, as laid down by the Supreme Court of the United States, which has raised many interesting questions at various times in United States tax history. There is no occasion here for dispute. We have simply to deal with two different concepts, the economic and the legal, and must be on our guard only against the danger of confusion from the use of the same technical terms in two different senses.

The purpose of taxation. Of course the primary purpose of taxation is the raising of money whereby the government may pay the cost of performing its primary functions. While the discussion thus far has ignored any other purpose, it is essential to recognize that taxation may be imposed for other purposes and to examine the economic bearing of the more important of these purposes. The taxing power is among the most powerful and far-reaching of the attributes of sovereignty. Even when applied only for the purpose of securing government income, its indirect economic effects may be — indeed certainly will be — very great. When consciously used for the accomplishment of other ends its power can

scarcely be exaggerated. Taxation may be used to impoverish one individual and enrich another. It may foster or kill whole industries and destroy the prosperity of entire regions. It may divert the course of industry and trade, affect the movement and location of population, interfere with the distribution and consumption of wealth. With such a power in its hands, it is not surprising to find the government occasionally using it for the accomplishment of ulterior ends, nor is it strange that numerous programs of social reform through taxation are constantly being urged.

The most familiar illustration is the protective tariff. The tariff, or customs, is a tax, imposed originally and primarily for the sake of revenue At various times and places it has been used for the purpose of checking the import of certain foreign products in order to relieve home producers of foreign competition and so to build up the home industries. This purpose has been so very prominent in the United States during most of its history that there may be a question whether protection is not really the primary purpose of the American tariff. Another example is the ten per cent tax imposed by the United States upon the notes of state banks. This tax was introduced with the sole purpose of legislating the state bank circulation out of existence. Revenue was not expected and has not appeared, since the tax has been successful in accomplishing its real purpose. Heavy license taxes imposed upon dealers selling alcoholic liquors are partly for the sake of revenue but also for the purpose of restricting and regulating the retail liquor establishments. Among suggestions for the use of taxation for "reform" are the proposal to limit or abolish roadside billboards by heavy taxes, the demand that inheritance taxes be used to produce a more equal distribution of wealth, and the time-honored "single tax" of Henry George.

In general it may be said that the use of a tax for another purpose tends to impair its effectiveness as a revenue machine. This is obviously true to the limit of such a tax as the prohibitory levy upon state bank notes, which produces no revenue at all, and it is especially applicable to the protective tariff. This subject will be pursued further in connection with our study of particular taxes, especially the tariff.

XXX

THE AMERICAN PROPERTY TAX

Evolution of the property tax. The possession of property furnishes so obvious a basis for the distribution of the costs of government or so obviously invites the exactions of the necessitous ruler, that it is not surprising to find that the property tax is among the oldest of taxes; history shows us numerous examples in ancient and medieval times. In those days and indeed for some centuries thereafter there did not exist the bewildering variety of different kinds of wealth and property rights which surround us at the present. Wealth consisted principally of lands and buildings; next in importance came domestic animals and agricultural tools and equipment; household furniture, clothing, and other forms of personal property were scarce and crude and, except in the homes of the kings and nobles, of little value. The earliest property taxes were generally imposed upon land and buildings. Later attempts to include various other forms of wealth and property rights as they gradually evolved did not prove successful, and the property tax in most European countries long ago returned essentially to its original base — real estate.

In America, on the contrary, where the property tax has since colonial days been the chief mainstay of state and local finance, the attempt is still generally made to include in its base most or at least very many of the complex modern forms of wealth and property rights. When thus inclusive, the tax is known as the *general property tax*.

Typical organization and operation in America. Being a subject of state law, the property tax is not carried out in exactly the same way in any two of the forty-eight states. However the general features of the tax present enough similarity to make it possible to construct a composite picture which will represent, not exactly the situation in any particular state, but rather the typical pro-

cedure which, with appropriate modifications, will conform closely enough to the situation in most of the states. The general property tax is provided by statute for the purpose of enabling the states and the counties, cities, towns or townships, and other local jurisdictions to secure the revenues necessary to their needs. The law determines what classes of property shall be taxable and what shall be exempt and prescribes a mass of administrative details. About the only legislative discretion left to the subordinate bodies is the power to fix the tax rates which shall be applied within their respective jurisdictions. On the other hand, the selection of assessors, collectors, and other administrative officers and the actual work of administration are almost wholly in the hands of either the towns or the counties, although there is a certain power of supervision and control lodged in the state tax commissioner or similar officer or board, the extent of which control varies greatly among the states.

Assessment. The first step in the operation of the property tax is the *assessment*, which means the discovery and valuation of all the taxable property of each person liable to the tax, as of a certain date each year specified in the statute. In some of the states assessment is the function of the towns; in the others it is performed by the counties. In this description we shall assume that the town has the matter in charge. The assessor must make out a "tax list" for each taxpayer, showing in more or less detail the description and value of all his taxable property. In some states the law requires the taxpayer to make out his own list and hand it to the assessor. In other states no list is required of the taxpayer, while in still others the taxpayer may be required to give certain information but not all, or he may have the option of filing a list or of having the list made by the assessor with an addition to the amount assessed as a penalty for failure to file the list. In every case however the final assessment is the work of the assessor, who is responsible for determining the total value of property to be assessed to each taxpayer. The law generally gives the assessor ample power to demand all necessary information from the taxpayer or from other witnesses.

Appeal, revision, and equalization. In each town there is a legally constituted body, known as the "board of relief" or by similar title

whose function is to hear and settle appeals from any taxpayer who may feel that his property has been wrongly assessed. When this board has finished its work the tax lists, as corrected, are combined to form the "town list," which contains at the least a list of the names of all owners of taxable property in the town with the total value of the property assessed against each.

Each town then reports the total amount of its tax list to a county board, commonly called the "board of equalization." The principal duty of this board is to determine after investigation whether the assessment in the several towns in its jurisdiction has been performed according to law, particularly whether the total value of property as assessed is equal to the full true value in each town. The board may arbitrarily increase or decrease the valuation of any town in order to make it equal to the full value, or at least to make the ratio between assessed and true value the same in all the towns. The chief reason for this equalization is that, as will presently appear, the county tax is to be apportioned among the towns in proportion to their respective tax lists, and any town which undervalues its taxable property or, more precisely, undervalues it in greater degree than the average of all the towns is thereby escaping a part of its share of the county tax and correspondingly burdening the other towns. Sometimes, though not usually, the county board of equalization may have authority to revise individual assessments.

Having corrected the lists of such towns as in its opinion require revision, the county board makes up its county tax list, which is a list of all the towns with the total assessed value of the property in each town. The total for the whole county is then reported to the "state board of equalization" or corresponding board or official. This board performs with respect to the several counties the same process of equalization as has just been performed by each county board with respect to its towns. The purpose also is similar; namely, to prevent evasion of part of its share of the state tax by any county through undervaluation of its taxable property. The final result, after the state board has made its revisions, is the state tax list, showing the total assessed value of property in the entire state as distributed among the several counties.

Apportionment. The tax rate. The next step is the *apportionment* of the tax. Each jurisdiction decides what amount of revenue it must obtain from the property tax; this amount divided by the total of the taxable property within its jurisdiction gives the *tax rate* for that particular unit. The tax rate applicable to any given taxpayer will be the sum of the rates of all the jurisdictions by which he is taxable To make this concrete, let us assume an example· Suppose that the work of assessment in a certain state has been completed, showing a tax list for the whole state of $1,000,000,000, which is distributed among the several counties as follows: County A, $200,000,000, County B, $100,000,000, and the rest in the other counties. Consider now County A, whose tax list of $200,000,000 is the sum of the lists of Town *a*, $50,000,000, Town *b*, $5,000,000, and the several other towns making up the remaining $145,000,000. Let us take account also of one other town, Town *c*, located in County B and having a tax list of $10,000,000 which is of course included in the $100,000,000 list of that county. Now let it be assumed that the state decides to raise $2,000,000 from the general property tax; the necessary tax rate will be calculated by dividing this amount by the state tax list, thus: $\frac{2,000,000}{1,000,000,000} = 0.002$, which in ordinary tax parlance would be called "2 mills on the dollar" or simply " 2 mills." In exactly this way each jurisdiction decides how much it must raise from the general property tax and then calculates its tax rate. Let the following be the assumed tax lists and amounts to be raised, with the resulting tax rates:

Jurisdiction	*Tax list*	*Amount to be raised*	*Tax rate*
State	$1,000,000,000	$2,000,000	2 mills
County A	200,000,000	200,000	1 mill
Town *a*	50,000 000	500,000	10 mills
Town *b*	5,000,000	100,000	20 mills
County B	100,000,000	200,000	2 mills
Town *c*	10,000,000	150,000	15 mills

We can now get back to the individual taxpayer. Suppose three individuals: M, living in Town *a*, N, in Town *b*, and P, in Town *c*, each having property assessed at $10,000. From M there will be collected the taxes of the state, County A, and Town *a*, to whom he must pay respectively 2 mills, 1 mill, and 10 mills upon each

dollar of his assessed valuation. His total tax will then be at the rate of 13 mills and will amount to $130. N will pay at the rate of 23 mills, being the sum of the rates of the state, County A, and Town *b*; his tax will be $230. P has a rate of 19 mills, made up of the rates of the state, County B, and Town *c*, his tax being $190.

Generally there will be other political bodies which, while having little or nothing to do with the administration of the tax, are permitted by law to collect a tax from the taxpayers within their geographical limits. Thus a borough is generally a part of a town which, being more thickly settled and having need for special town services not performed for the whole town, is permitted to vote for such special services, to be paid for by a borough tax at a sufficient rate as determined by the people of the borough or their proper officers. If for example the individual M in our previous illustration lived in such a borough within the Town *a* and the borough tax rate were 7 mills, his total tax would be at the rate of 20 mills and would amount to $200, whereas another resident of the town living outside the borough would pay only the 13 mill rate The same situation arises in the case of cities, whenever the city is not coterminous with the town. The school districts are the most important and numerous of the minor taxing jurisdictions; they very commonly have independent authority to determine their school revenues and their district school tax rates. There are also fire districts, street lighting districts, road districts, and other examples of limited groups of the people who are organized for the sake of obtaining special government services with authority to tax themselves by adding their own rates to the general property tax.

Collection. The final act is the *collection* of the tax. Each taxpayer receives a tax bill, or formal notice of the amount of his total assessment, the tax rate, and the amount of the tax due from him. In some states each town and county and the state has its own collector, who sends out bills and collects the tax due. The taxpayer may thus receive two, three, or more separate tax bills and make as many separate payments. In other states one bill is sent out and the entire tax is collected by the town or county collector, who then distributes its proper share to each jurisdiction concerned.

Modifications. The reader should again be reminded that the foregoing description is intended to be typical rather than descriptive of any particular state, and that a great number and variety of modifications is to be found. Equalization is in many states not performed by both county and state boards. There are certain states in which the state government has given up all share in the proceeds of the property tax and so has little or no interest in equalization. The process of apportionment is not always carried out in all the refinement here described, and in at least one state an entirely different method is employed. There are states in which certain of the tax rates are fixed by statute in advance of assessment, so that the assessment determines the amount of the yield. Such modifications in detail, of which these are only a few examples, are numerous enough but need not concern us here.

Position of the property tax. The tax whose practical machinery we have thus examined at some length is the main source of revenue of the states and local bodies of the United States. In the year 1932 according to the United States census [1] it produced for all the states and their subdivisions 4,685 million dollars, 53 per cent of their total revenue receipts of 8,852 millions The reliance of the states and the several subordinate grades of government upon the property tax is clearly shown in the following table [2]

	Total revenue receipts	*Property tax receipts*	
		Amount	Per cent of Total
State governments	$2,208 mil	$ 323 mil	15
Counties	1,314 mil	877 mil	67
Cities, towns, villages, and boroughs	3,374 mil	2,057 mil	61
School districts	1,515 mil	1,119 mil	74
Townships	248 mil	195 mil	79
Other civil divisions	194 mil	112 mil	58
Total	$8,852 mil	$4,685 mil	53

[1] *Financial Statistics of State and Local Governments, 1932*, 1935, p 7

[2] *Ibid*, pp 4 and 7 See also the following reports of the National Industrial Conference Board · *Taxation and National Income*, 1922, *Tax Burdens and Exemptions*, 1923; *Tax Burdens and Public Expenditures*, 1926, *Cost of Government in the United States, 1929–1930*, 1932, and *Cost of Government, 1923–1934*, 1934; also reports of the United States Bureau of the Census upon financial statistics of states and of cities.

The total tax bill of the American people (to all grades of government, federal, state, and local) in 1938 was 14,811 million dollars. 6,034 millions of this were collected by the federal government. According to estimates of the United States Treasury Department, issued in August, 1939, the property tax in 1938 produced 4,745 million dollars for all the state and local governments. This was equal to four-fifths of the total tax collections of the federal government and accounted for one-third of the entire national tax bill (for all grades of government). Evidently the property tax is, on the basis of amount collected, by far the most important tax borne by the American people.

Definitions. This is an appropriate place to define certain technical terms commonly used in property tax discussions. *Real estate* means in general (1) land (in its broadest sense), (2) all buildings and other structures permanently attached to the land, and (3) such equipment, machinery, etc., as is so definitely attached to a building that it could not be removed to another location without serious impairment of its value or use, such as elevators, fire escapes, shafting for machinery, etc.

Personal estate or *personal property* or *personalty* means all possessions except real estate. Personal property is divided into two classes: (1) *tangible* personal property, being that which has a physical substance, such as machinery, household furniture, jewelry, etc., and (2) *intangible* personal property, which includes all other personalty; that is, abstract rights against wealth in the physical possession of others or against other persons, such as promissory notes, corporation stocks, bank deposits, etc. It will be noted that the word "property" is here somewhat loosely used in conflict with the economic definition of that term. Strictly speaking, intangible personal property alone deserves the title of "property"; the others – real estate and tangible personalty — are "wealth." In most discussions of taxation however the term "property" is broadened to include all physical wealth as well as abstract property rights, and this usage is so well-nigh universal that it is scarcely to be avoided. If we occasionally follow the popular usage, no harm will be done so long as the sense in which the word is used is clearly understood.

Assessment: Discovery of taxable property. The crucial point in the practical operation of any tax is the assessment. The general property tax requires that there be entered opposite the name of each person a list of his taxable possessions (which includes nearly everything that he owns), with a statement of the true market value of each item and of the total. This implies (1) the discovery of all such property by the assessor in the face of at least the passive resistance of the taxpayer, and (2) the process of valuation. As to the first, the task is easy enough in the case of land and buildings and the more bulky and less movable kinds of personal wealth, such as farm and factory machinery, farm animals, etc. But the task of locating other kinds of tangible personalty, such as jewelry, is far from easy. Even such property as household furniture, rugs, books, pictures, clothing, etc., can be discovered only by a personal search of the taxpayer's home, and although the tax assessor generally has the legal right — indeed the obligation — of such search, he is loath to resort to it in the face of the natural repugnance of the taxpayer. House-to-house search for personal property is today the exception, not the rule, in tax assessment. If it were the rule, it is safe to say that a considerable amount of the more valuable and easily movable articles would be spirited away and concealed on tax day.

When we come to intangible property, difficulty of discovery becomes practical impossibility. Bonds, notes, stocks, bank deposits, and similar classes of property are readily concealed in bank vaults or in the taxpayer's office or home or if necessary may be actually removed into another town or state over tax day. The assessor is powerless, and the best he can do is to accept such information as is volunteered by the more conscientious taxpayers or to make a pure guess at the amount of such property and trust to the taxpayer to protest in case he has overshot the mark. Intangible property such as real estate mortgages, which must be publicly recorded, is discoverable, but this is no help when the mortgaged property is located and the mortgage recorded in another jurisdiction from that of the taxpayer's residence and tax obligation, since the assessors can hardly be expected to search the records of other towns or counties. Moreover most states exempt mortgages

upon real estate within the state, thus removing the only ones that could easily be discovered.

Practical problems of valuation. The state laws differ as to language, but in general they require that all property be assessed at its true market value, or words to that effect, sometimes specifying that it must be the value at which the property would be sold by a willing seller to a willing buyer, seeking quite properly to exclude the evidence of forced or otherwise unrepresentative sales. Tax legislation thus agrees closely with the economic concept of value.

Now except at the very time of a sale, the value at which anything "would sell" — in short its *value* — is a matter of judgment. It must be determined by an appraisal, which is just what assessment really is. Anyone acquainted with practical affairs has an idea of what is meant by appraisal. He knows that it can be accomplished only by an expert in the particular kind of property involved who is in possession of all the pertinent facts. Land appraisals are made for business purposes by real estate experts, who first gather carefully all available evidence and then give ample time to the study of the problem. They are usually assured of the coöperation of the owner or of someone else practically interested in the matter. The appraisal of a large apartment house or office building or factory is a difficult technical matter, which only the specialized expert can handle. And here again, when we pass from real estate to personalty, we pass from the difficult to the more difficult. The appraisal of the machinery in a cotton mill, of the raw materials, the "goods in process," and the finished products of a brass factory, the stock in trade — "inventory" — of a dry-goods store is each a complicated, time-consuming task for the man especially equipped for that particular appraisal. If it is instead a silk factory or a hardware store, another set of experts will be required. Such appraisals are constantly being made in practical business, for insurance companies, banks, investors, brokers, and for the manufacturers and merchants themselves. To be accomplished with the accuracy demanded for business purposes, they require the careful labors of specialized technical experts, highly trained and experienced.

Theoretical questions. Furthermore there are theoretical problems of valuation which are not solved merely by employing the right experts. How for example shall we assess the partly finished "goods in process" of a manufacturing plant? The manufacturer has his own methods for inventory purposes, based upon more or less arbitrary accounting rules; he may take cost of raw materials, add direct labor cost, put in something for "overhead," etc. But, except as an accounting fiction, the "value" thus obtained means little; it is to be doubted if it is really value at all. Who would buy the partly finished goods of most factories? Certainly it is questionable whether a "value" thus obtained is the true market value prescribed by the tax law. And if this is not the legal tax value, what in the world is? Another accounting rule is "cost less depreciation." But if there is no one likely to want to buy the article, can we accept the results of such a "valuation"? What for example should be the method of assessing special tools or patterns made and used in a factory on a particular job which is now completed? These are merely a few samples of some of the theoretical and technical questions which may well puzzle even the expert appraiser or the lawyer.

The assessor and his job: Expert assessment. The foregoing description of the assessment process will suggest that here is a difficult technical task, one that would clearly appear to require for its successful accomplishment a personnel of trained expert appraisers equipped with adequate organization and staff.

These requirements are met, more or less adequately, in many of the larger cities and in some small cities. A staff of trained assessors, with the necessary clerical and other assistance, is built up. The work is so arranged as to continue through the year, permitting of long term or permanent employment, and giving to the assessor a steady "career" job, with adequate salary. Accurate maps of the entire region are made and skilfully used. Complete descriptive data regarding land and buildings are prepared and recorded. Evidence regarding all sales is gathered. Scientific rules for determining values with respect to location and various other standards are used. Even the aeroplane is pressed into service for the quick mapping of extended areas in the suburbs or country districts. In

many an American city, where such conditions prevail, a high degree of efficiency, in the assessment of land and buildings at least, has been attained.

Some of the states have obtained a certain degree of centralization of the work of assessment in the office of a state tax commissioner or board, which has enabled them to secure full-time services of engineers and other experts to aid in the work of local assessment.

The ordinary local assessment: General situation. But in spite of assistance now and then from the state tax commission, assessment is still essentially a local function, and the ordinary town or county or small city will seldom have anything like an adequate assessment organization or procedure. There is not enough work to warrant a full-time job with attractive salary. Expert qualifications cannot be and are not required of the candidate for the assessor's job. Let us picture for a moment the typical town assessor as he enters a great factory devoted to the manufacture, let us say, of clocks and watches, with millions of dollars' worth of land, water power, factory buildings, machinery and equipment, raw materials, goods in process, finished products, not to mention such intangible property as patents, contracts, "good-will," investments of various sorts, and so on An appraisal of this plant and property may have been recently made for the owners or for a syndicate of bankers interested, which took weeks or months of labor by a corps of experts and cost hundreds or thousands of dollars. The assessor, if not the only one in the town, is one of two or three or half a dozen. He is perhaps a farmer or a carpenter, he was never in a watch factory before, never saw or even heard of many of the kinds of wealth before him. This factory is simply one of hundreds or thousands of properties which he must assess in the space of two or three months at most He cannot take the time even to glance at all the items of wealth here assembled. Yet he is the one who is required by law to put a value upon this complicated assembly of wealth and property, which shall in large measure determine the amount of the contribution of its owners toward the cost of the town, county, and state governments.

It should be recognized that all assessments are not so complicated and so far above the heads of the assessors. The local assessor

in an agricultural town or a small village, himself a farmer or carpenter or blacksmith, gets on after a fashion in the assessment of farm lands, buildings, and equipment, and the simple homes and personal effects of the village. Even the stock in trade of the country store may be fairly within his ken, and he may not come up against the really difficult problems. But let no one delude himself with the thought that even this simple kind of assessment is as a rule well done; the facts are to the contrary. And it is the exceptional town which does not contain some examples of taxable property that will put to rout the equipment of any ordinary assessor.

Information from the owner. It may be asked: why not compel the owner to give the assessor the advantage of the results of his own appraisal or inventory or accounts; or does not the law in fact now give the assessor power to demand such information from the taxpayer? The answer is that, the tax assessor being unable to check the accuracy of such reports, reliance upon them would result in a system of self-assessment. Since even the most expert appraisal is a matter of judgment, and since there is the greatest variety in accounting practice, errors and inequalities would be inevitable, even assuming an honest intention on the taxpayer's part to make a truthful return On the other hand, with the strongest of motives and almost unlimited opportunity, the probability of false statements and dishonest collusion is too great to be ignored. American tax practice has never fully accepted the principle of self-assessment, and while a very considerable amount of information is required of the taxpayer by the law, the ultimate obligation of the assessor to view and value the taxable property is seldom waived Furthermore the very natural disinclination of the taxpayer to disclose to any outsider the intimate details of his personal possessions or the technical facts of his business is so strong that it has generally resisted successfully the assessor's efforts to secure complete information.

The attitude of the taxpayer. Under such conditions and with human nature as it is, the assessor can hardly expect much help from the taxpayers. There are those indeed — and their number is surprisingly large — who conscientiously try to pay all that the law

requires, even though there be little to compel it. Such persons are rewarded for their extreme honesty by bearing a burden of taxation out of all proportion to that borne by the rest of the community. The majority take the attitude of passive resistance: "Let's see you get it if you can." And a few use every means, fair and foul, to evade as much of their tax burden as is possible. So general is the cynical acceptance of this situation that the majority of honest, law-abiding citizens regularly once a year take oath to tax lists which they know perfectly are false on their face, and if any taxpayer was ever prosecuted for this open perjury we have not yet heard of the case.

How the assessor meets the situation. Under such circumstances as these, we need not be greatly surprised at the result. The assessor generally does about the best he can. In the first place, he has last year's lists, of which he takes full advantage; copying last year's figures will suffice for the majority of cases where he has reason to believe there has been no material change. He will travel over his district, making brief inspections of the outside of new buildings, casting cursory glances at the merchants' shelves, asking a few questions, picking up what gossip he can as to the worldly goods of his fellow townsmen. Using such information as this, reinforced by a good deal of often shrewd guesswork, influenced almost always by the popular satisfaction which will follow when everybody feels that he has been assessed considerably under true value, influenced now and then no doubt by the desire to be kind to a friend and to stand well with those of political and social influence, seldom downright corrupt, the assessor manages somehow to check up on the lists handed in by taxpayers, to make out the lists of those who fail to file them, and finally to turn in a town tax list which will not cause too many appeals to the board of relief, nor be too much censured by the county or state board of equalization, nor cause so much local dissatisfaction as to threaten his reëlection.

Results of assessment. With all the circumstances as thus pictured, we should hardly expect to find a true and faithful assessment of all taxable property, and the facts will fully justify our surmise. For reasons which are now obvious, the assessment of real estate is more effective than that of personal property. Yet figures collected

by the United States Census, going as far back as 1890, show that the assessed value of real estate all over the country is not more than half the true value legally taxable. These figures are confirmed by investigations by state tax commissions in various states and by numerous tax students. And while on the average real estate gets onto the tax lists at not more than half its true amount and value, this gives no adequate picture of the extreme variation among individual properties, which are assessed all the way from zero (when the property is not on the tax list at all) to 100 per cent of the value or even more. A careful investigation made in 1922 by the Connecticut State Tax Commissioner [1] showed in one small Connecticut town no item of real estate assessed at more than 40 per cent of its estimated true value, while the lowest was 9.2 per cent, and the average of all the cases investigated was 24 per cent. In another town, somewhat larger, the percentages of assessed to true value ranged all the way from 41 per cent to 106 per cent and the average for the thirty cases investigated was 76 per cent. Note not only the extraordinary variation among particular assessments but the great discrepancy between the two towns. For the whole state the average ratio of assessed to true value of real estate was probably not over 60 per cent. There is similar evidence from other states and no reason to believe that any state would show conditions much better.

When we examine the results of similar investigations of the assessment of personal property, the figures are either appalling or ridiculous according to one's state of mind. The United States Census, beginning with 1890, finds that less than 20 per cent of the true value of taxable personalty gets on the tax lists. As far back as 1870 a New York state commission estimated the ratio at 15 per cent in New York. There is every reason to believe that the breakdown is becoming more complete from year to year. A special investigation conducted by the Connecticut State Chamber of Commerce in 1916 [2] showed the following results as regards the assessment of personal property upon the farms of the state: only

[1] *Report of the Tax Commissioner* (of Connecticut), for 1921 and 1922, pp. 6-16.

[2] *Report of the Joint Committee on Taxation and State Finance to the State Chamber of Commerce*, 1917, pp. 12-18.

55 per cent of the true value of cattle was listed for taxation, 17 per cent of the cattle being not assessed at all; of sheep, goats, swine, and poultry less than 24 per cent of the true value was assessed; 41 per cent of the true value of horses and mules was reached; the assessed value of farm machinery was 6 per cent of its true value. And it must be remembered that these are among the classes of personal property most readily discoverable and most easy to value. Investigations into the assessment of household furniture, books and libraries, musical instruments, watches and jewelry, and other such classes of personal property demonstrate the almost complete escape of such possessions. And finally, when we come to intangible personal property, the value actually assessed is nowhere more than a trifling fraction of the amount of such property in existence and legally taxable. It is no exaggeration to say that the escape of intangible personal property throughout the United States is almost complete and that in most parts of the country the assessment of important classes of tangible personalty is either a ridiculous failure or has been virtually abandoned.

Undervaluation and high tax rates. Certain collateral results of imperfect assessment are of special importance. Consider the relation between prevailing undervaluation and property tax rates. Assessed values are almost everywhere below true values; sometimes by a moderate margin, sometimes so far that they are only a small fraction of the real value. Assessments at 20 per cent, 15 per cent, 10 per cent of true value or even less are no curiosities. There are two reasons. First is the natural desire of almost every taxpayer to get his own property in at the lowest possible value, in order that he may gain at the expense of his neighbors. In the second place is the common interest of the whole town in a low scale of assessments, since thereby the town reduces its share of the apportioned county and state taxes. Each taxpayer is therefore glad to have all assessments low and would in addition like to see his own at least a little lower than the average. Such universal pressure is too much for the average assessor, generally owing his position to the votes of the taxpayers.

Of course as certain classes of taxable property escape in whole or in part from the general property tax, either through failure to be

listed at all or through undervaluation, the burden upon other forms of wealth which are more effectively reached is correspondingly increased. The states, counties, and towns must have their necessary revenues, and as the tax base is narrowed by concealment and undervaluation, the tax rates have to rise. Tax rates have been going up steadily all over the United States, virtually since the beginning. Rates of 20 mills, 30 mills, 40 mills are common; 80 and 90 mill rates are not unknown. Such rates mean confiscation except when tempered by undervaluation. Undervaluation is the rule for most kinds of tangible wealth, real and personal, but for many kinds of intangible property undervaluation is hardly possible. If the assessor discovers a $10,000 mortgage at a good rate of interest upon adequate security he can scarcely enter it on the tax list at any other value with that figure staring him in the face. Corporation stocks and bonds, regularly quoted in the market reports, can hardly be entered at anything but their true market values. Now a tax at the rate of 2 per cent to 4 per cent upon a 4 per cent bond means anywhere from taking half the value to complete confiscation. This will be resisted by any intelligent taxpayer whose conscience is not quite quixotic and will indeed make the assessor himself hestitate; intangible property is virtually forced whenever possible to escape by means of concealment. We have a vicious circle; the higher the rates the more complete will be the evasion of all property which can possibly escape, and the more property succeeds in escaping the higher will the rates have to be on that which is left. Such property as cannot escape assessment is able to endure prevailing tax rates only because of prevailing undervaluation, and thus low valuations and high tax rates pursue each other about another vicious circle. Real estate has thus come to bear the bulk of the tax burden; at least four-fifths — probably more — of the property assessed consists of real estate.

Inequality and the legal remedy. With complete escape of certain kinds of property the regular thing, with gross undervaluation the general rule for all property that cannot escape, with tax rates so high as to be unendurable except when accompanied by undervaluation, with variation in individual assessments all the way from complete escape to more than 100 per cent of true value, necessity

of demonstrating serious inequality among the several taxpayers scarcely exists. We can safely let it go at that, and proceed to inquire as to the operation of the legal machinery for relief.

Any taxpayer who feels himself unjustly treated in the assessment of his own property may present his complaint before the local board of review. But except in the few cases where there is actual assessment at more than the true value, this generally avails him nothing. If in a town where property is assessed on the average at 40 per cent of true value, a certain taxpayer is assessed at 80 per cent, it is obvious enough that he is unjustly treated. But all he can prove to the board of relief is that his property, even so, is assessed 20 per cent less than the law requires. There have been a few cases in which the courts have required reduction of a particular assessment to agree with the general practice, but since proof of a general practice usually requires bringing evidence as to the true value of taxable property of every other taxpayer, or the majority of taxpayers, this remedy is seldom effective. Justice by legal review can be fully effective only when the millennium of full value assessment shall have been reached.

The attempt to give administrative relief for wholesale inequality caused by general undervaluation by towns and counties is equally ineffective. The county board of equalization has usually no possible way of determining the true value of taxable property in each of the towns. The most it can do is to gather evidence as to a few particular assessments, listen to the gossip that is current, and mainly by guesswork raise this town and lower that in the hope of diminishing some of the most glaring inequalities. The review of the state board of equalization is equally ineffective, though there are a few states in which a competent tax commission, armed with authority to revise original assessments in the towns, has been able to accomplish some good.

Confusing wealth and property. Some of the most serious difficulties of the property tax arise out of a fundamental defect in the very nature of the general property tax idea; namely, confusion between the concepts of wealth and property. If all physical wealth were taxed and in addition all the property rights to wealth, there would be no injustice, since wealth and property rights to wealth

are coterminous; there would simply be an unnecessary duplication of taxes where one tax, on either wealth or property at double the rate, would accomplish the same result. But many forms of property have no legal document to represent them and hence do not come within the purview of the tax laws. The result is that our laws usually attempt to tax *all* forms of physical wealth (barring the special exemptions) and in addition *some* classes of property rights. For example A and B are merchants, each owning a store and fixtures and stock in trade assessed at $100,000. A incorporates his business, the entire capital stock being owned by himself and members of his family. Now if, as many states require, his corporation is taxed upon the store and fixtures and goods assessed at $100,000, while he is taxed upon $100,000 of capital stock in the corporation, his tax is twice as much as it was before and twice as much as B's tax, with absolutely no difference in taxpaying ability. This is double taxation of the crudest sort. The same thing results when the holder of a mortgage or promissory note or other credit is taxed upon it, while the debtor is taxed upon the entire value of his possessions without deduction of the amount of the debt.

The present attitude of American state tax laws on this matter may be roughly summarized as follows: (1) shares of stock of domestic corporations are taxable to resident owners in somewhat less than half the states; shares of foreign corporations are similarly taxable in somewhat more than half the states; (2) bonds of corporations, whether domestic or foreign, are taxable to resident owners in about three-fourths of the states; (3) mortgages are generally exempt; when secured by property within the state, nearly all the states grant exemption, and a large majority exempt even mortgages secured by outside property.[1]

Modifications in detail. Attention should be called to certain adjustments in detail which have served somewhat to improve the operation of the general property tax. In some states the state has withdrawn from participation in the revenue from the tax, thus removing one cause of undervaluation. But if county apportionment continues, the improvement is not great, and there may be

[1] Cf. National Industrial Conference Board, *State and Local Taxation of Property*, 1930.

actual loss through withdrawal of state supervision and control. The state of Connecticut has solved this problem in the one effective way by apportioning the state tax, not upon the assessed value of property in the towns, but upon the respective amounts raised by the property tax.[1] Since reduced assessments must be accompanied by higher local tax rates in order that the town may secure its own revenue (the product of assessment by tax rate), the town has nothing to gain by undervaluation so far as the state tax upon the town is concerned.

Recognition of the peculiar position of intangible property, which has been pointed out, has led a number of states to provide for the taxation of such property at rates much lower than those applied to tangible wealth. It has been the hope that the taxpayers when granted a reasonable rate would voluntarily list their intangibles, and the experience of several states has shown a material increase in the amount of such property assessed following a reduction in the rate. There is no evidence however to show that any state has by this expedient succeeded in assessing the whole or even the greater part of the intangible property subject to its tax laws, and there is considerable evidence to show that the low-rate remedy gives at best only a slight mitigation of the former evasion.

Reform of the property tax. Adequate treatment of the topic indicated by this title cannot be compressed within the limits of a general text such as the present. A few suggestions only will be offered. Reform of the *general* property tax is not possible; only by relinquishing the idea of a general property tax can relief be expected. The first step is certainly the abandonment of all taxation of intangible personal property. This is indicated both by the theoretical nature of intangible property rights and by the complete and inevitable breakdown in administration This step has been recommended by numerous tax authorities, including a committee of the National Tax Association, and it has already been accomplished by some ten states.

Next in order is the abandonment of the attempt to tax all those classes of tangible personalty where experience has shown the prac-

[1] Connecticut has unfortunately continued to reward undervaluation by certain grants of state moneys to the towns with low tax lists.

tical impossibility of a real assessment. This would include household furniture, private libraries, musical instruments, clothing and jewelry, and similar objects of essentially personal or domestic character. We may as well admit that these classes of property never have been and never will be properly assessed and give up attempting the impossible.

A special problem is presented by machinery and goods and materials of manufacturers and merchants. We have the choice of either giving up the taxation of such goods, as has been done by New York, or consigning their assessment to the state tax commission, with authority to use when necessary specified rules of assessment which do not necessarily arrive at market value. For example it would frequently be practicable to use methods such as are employed by the professional accountants in valuing these classes of business equipment and materials, basing estimates upon cost of replacement less depreciation, or upon cost of raw materials plus cost of labor in fabricating them. Safeguards should be provided however in order to prevent too wide departure from the fundamental standard of the value of the taxed wealth. This second remedy has recently been recommended by a special tax commission in Connecticut.[1]

These changes as to personal property would leave for assessment by the local assessors only real estate and a few forms of bulky, easily discovered, and easily valued tangible personal property. What should be substituted in order to reach the taxpaying ability possibly represented by the classes of tangible property exempted is not so serious a question as might at first appear. The fact is that, since the classes of property which it is proposed to exempt are the ones which now generally escape and which yield comparatively little revenue, the loss of revenue to be made up from this reform is not great. No substitute for household furniture, etc., is necessary. The amount of such property bears a fairly close relation to the value of the homes, and the microscopic increase in the tax rate on real estate that would suffice to make good the lost revenue would cause no injustice. As to industrial machinery and goods and

[1] *Report of the Connecticut Temporary Commission to Study the Tax Laws of the State and to Make Recommendations concerning Their Revision*, 1934, Chap. X.

materials of manufacturers and merchants, in case such wealth is to be removed from the property tax, some form of business income tax offers a practicable substitute. Thus New York increased her tax on the net income of business corporations in lieu of the tax on their personal property.

We must still take account of the fact that the present assessment even of real estate is far from perfect. Here however the task is not inherently impossible, and there is reasonable hope for improvement, once the assessors are relieved of the obligation to do the impossible and are allowed to concentrate their efforts upon that which can be done. As already suggested, there are certain cities whose assessment of lands and buildings has attained a high degree of efficiency. Similar results are possible in any city, town, or county where the finances are on a large enough scale to warrant the expense, and for the smaller and less wealthy towns and counties a consolidation of the work of assessment in the hands of larger units or groups of units is possible. While the work of assessing real estate may never be so organized as to produce perfect results, there is no reason why any state should not provide assessment machinery accurate enough to accomplish its purpose; *i.e.*, the equitable distribution of the cost of government.

A certain degree of centralization of administrative authority in the hands of a state officer or board is a great help in the task of assessment. This has been obtained in some states, with beneficial results. It makes possible the employment of technical experts whom most towns and counties cannot afford. Statistical data, rules of appraisal, and assistance of various sorts may be furnished the local assessors, and the authority of the state may be exerted to compel more efficient work on the part of local officials.

Any reform of the property tax requires that assessment be on the basis of the full market value of all taxable wealth and that tax rates be so low as to call for a reasonable, not a confiscatory, contribution from the owners. These two features go hand in hand. Assessed values cannot be brought up to 100 per cent if present high tax rates are to continue, and present rates cannot be materially reduced without increase of the tax base; on the other hand, either reform becomes possible if accompanied by the other. The

motive to undervaluation through the present method of apportioning state and county taxes must be abandoned; the substitute of the Connecticut method based on the tax revenue is a simple way out. When once the habit of full value assessment becomes established, review by a board of relief may be made to mean something, and an aggrieved taxpayer may have some real remedy.

Taxation of corporations. The taxation of corporations presents several special problems. In the first place, there is the complex problem of double taxation resulting from taxing both the corporations on their property and the owners on their stocks and bonds, to which reference has been made. Secondly it is the corporations which are the principal owners of those complicated categories of property whose valuation proves too much for the ordinary town or county assessors. In the third place, the wealth of many great corporations is located in more than one taxing district, being spread over several towns or counties or even states. The property cannot be assessed piecemeal, and no assessor has authority to assess the whole. This situation is peculiarly evident in the case of the railroads, telephone and telegraph companies, and other public service companies. These problems have been met in many states by taking the assessment from the local bodies and entrusting it to a state commission. In addition special methods are frequently applied, very generally to the public utility corporations and the banks and less commonly to the general business corporations.

Conclusion. The property tax, with all its defects, is still the backbone of the American tax system. In spite of the development of various other forms of taxation, which have permitted many of the states to reduce their reliance upon this tax, the property tax still contributes approximately half of the total tax revenue of the nation, including all grades of government. Indeed it is the property tax which, during the years of depression since 1929, has proved the anchor to windward that has saved American public finance from a vastly greater shipwreck than it has actually suffered. The relative contribution of the property tax, as compared with other taxes, has materially increased during the depression.

On the other hand, the property tax, as it exists in most of the states, is characterized by glaring defects which cause serious

inequality and injustice in the distribution of its burden. These defects have been sufficient to lead many students to desire and predict the decline and final passing of the property tax. Such predictions have not been fulfilled, and are not likely to be. The property tax has maintained itself because of its absolutely essential contribution to the cost of government. This however should not blind our eyes to the injustice which it perpetrates. Fortunately the defects of the property tax are not incapable of remedy, along lines which have been pointed out. Progress is moving, with some considerable rapidity, in this direction. America is probably destined to follow the European example in giving up the fiction of a *general* property tax. Thereafter reforms which are practicable and readily discernible should make of the property tax, not merely an effective revenue producer, but also a defensible instrument for the equitable apportionment of the costs of government in the American states.

Shifting and incidence. The problems of the shifting and incidence of the property tax are so closely interwoven with the corresponding problems of the income tax that their discussion will be most conveniently taken up after our study of the income tax. This topic is therefore postponed to the close of the next chapter.

XXXI

THE INCOME TAX AND OTHER TAXES

The income tax: Historical background. Whereas the property tax is one of the oldest forms of taxation, the income tax is a modern tax. The first real income tax was introduced in England in the year 1798. It was intended as only a temporary measure, though it actually lasted until 1816. It was not until 1842 that England again turned to the income tax, but from that year this form of taxation has been a permanent part of the British revenue system. Italy adopted the income tax in 1864. In Germany the separate states introduced the income tax at various dates during the latter half of the nineteenth century. Japan fell into line in 1887. France did not make use of the income tax until just before the outbreak of the World War. In Germany the income tax was taken from the separate states and made a national tax in connection with the revolution after the World War. From its beginning the income tax has won recognition as an acceptable form of taxation, and it has grown steadily in importance, until today it is the backbone of the national tax systems in most of the leading nations of the world. In the United States the income tax was used by a considerable number of the states before the federal government took its first step in this field. Some state income tax laws were enacted between 1840 and 1850, and during the second half of the nineteenth century the number grew to sixteen. These laws all carried over enough of the administrative features of the general property tax to insure their complete failure. Assessment was based upon the personal declaration of the taxpayer, and administration was in the hands of local officers; by the end of the nineteenth century the breakdown was complete.

The federal government first made use of the income tax at the time of the Civil War. This was a war measure, its yield declined rapidly after the war, and it was repealed in 1872. In the year 1894 the national government made its second excursion into the

field of income taxation. This act was declared unconstitutional before it went into effect, upon grounds which we shall examine presently. It was not until 1913, following an amendment of the federal Constitution, that the income tax became a permanent part of the tax system of the United States government. About this time state income taxation also acquired a fresh lease on life, beginning with the Wisconsin law of 1911, a carefully drawn law providing for complete administration by state officials. It has been a distinct success. Since then some thirty states have enacted personal income tax laws, which, profiting by the experience of the earlier epoch, generally provide for such centralized state administration as to give them at least a fair chance of success. Among the more conspicuous examples are those of New York and Massachusetts. A few states also levy income taxes upon corporations only. Today the income tax is the principal tax of the federal government and a source of revenue, more or less important, in more than half of the states.

Income taxation and the United States Constitution: Before the sixteenth amendment. The federal income tax of 1894, though never in effect, raised constitutional questions which are of great importance to the student of American taxation. Congress is forbidden by the Constitution to levy any direct tax unless apportioned among the states according to their population. This prohibition is expressed in Article I, section 9, in the following language: "No capitation, or other direct, tax shall be laid, unless in proportion to the census or enumeration hereinbefore directed to be taken." This means for example that, if the population of state A is double that of state B, the contribution of the residents of state A under any direct tax must be double the amount contributed by the residents of state B. This proportion must hold regardless of the wealth or incomes of the people. It might happen, for example, that B was a wealthy state and A a poor state, so that while A's population was double B's, the wealth of its residents was half that of the residents of B. Then, to carry out the constitutional rule of apportionment, the rate of the tax on the wealth of the residents of state A would have to be four times the rate imposed in state B. This rule is so flagrantly unjust that Congress has never attempted to

collect any large amount by means of direct taxes, and since the Civil War no attempt to use direct taxes has been made. The constitutional rule is practically equivalent to a prohibition of direct taxation.

The opponents of the income tax of 1894 claimed that the income tax was a direct tax and that, since the law made no provision for its apportionment according to the constitutional rule, it was therefore unconstitutional. Now the income tax, as regards its economic character, is certainly a direct tax, in that its burden is normally borne by those on whom it is imposed and not shifted to others. But this was not the idea of a direct tax in the minds of the framers of the Constitution, who apparently were thinking only of land taxes and poll taxes. The income tax was accordingly considered an indirect tax, legally, and it was so held by the Supreme Court of the United States when this question was raised with regard to the Civil War income tax. In 1894 however the Court reversed its previous ruling and in the case of *Pollock v. Farmers' Loan and Trust Co.* (157 U. S. 429) held that the income tax was a direct tax, thus bringing the legal definition into practical harmony with the economic concept of direct taxation, though by means of a line of argument which will scarcely stand the test of economic analysis. Being a direct tax, the income tax of 1894 was of course unconstitutional since it was not apportioned among the states according to population. Such an apportionment, as we have seen, would have been intolerable. This decision thus practically deprived the federal government of the power to use the income tax.

The Sixteenth Amendment. However the popular demand for this form of taxation was so strong that the constitutional obstacle was not permitted much longer to continue. In 1909 Congress imposed a tax upon the incomes of corporations only, this tax by a twist of legal reasoning being regarded as an excise—hence an indirect tax—and so not requiring distribution according to the population of the states. At the same time Congress proposed an amendment to the Constitution, which was finally ratified by the requisite number of states and promulgated as of March 1, 1913. This amendment, without reopening the question of definition or altering the constitutional rule of apportionment for other direct

taxes, expressly gives to Congress the power to impose an income tax without the necessity of apportionment. The ratification of this amendment was immediately followed by the enactment of the income tax of 1913, which imposed a tax upon the incomes of individuals and corporations, absorbing the existing corporate income tax of 1909.

The present United States income tax:[1] **Individual incomes.** The law distinguishes between individuals, corporations, and partnerships. As to the first, all residents of the United States are in general subject to the tax upon all their taxable incomes. Nonresidents are generally taxable upon income derived from sources in the United States. The tax is upon the *net income*, which as defined in the law is determined in the following manner: The taxpayer's *gross income* consists of all receipts except those specifically excluded by the law. The principal categories thus excluded are the proceeds of life insurance policies, gifts, bequests, interest on state and municipal obligations, and (to a certain limited extent) interest on United States bonds and notes. Items so excluded need not be reported at all in the taxpayer's return.[2] From the gross income the law permits certain *deductions*, which include the expenses necessary to obtain the gross income, together with taxes, interest, and certain losses and contributions. What remains after subtracting the deductions from the gross income is the *net income*.

The individual tax is in two parts, a *normal tax* and a *surtax*. Before calculating either tax, the taxpayer is permitted to subtract from his net income two *credits:* (1) a "personal exemption" of $1,000 for a single person or $2,500 for a married person or a single person who is head of a family (the combined personal exemption of a husband and wife may not exceed $2,500, which amount may be taken by either one or divided between them), and (2) a credit of $400 for each child under 18 years of age or each person incapable of self-support who is supported by the taxpayer. The remainder obtained by subtracting these credits from the net income is the "surtax net income."

[1] Federal revenue act of 1938

[2] Certain tax-free interest is reported for information only.

To obtain the base subject to the normal tax, we start with the "surtax net income" and deduct therefrom the following two credits: (1) interest on such United States bonds as were included in the gross income, and interest on obligations of certain corporations organized as instrumentalities of the United States government (such as the Reconstruction Finance Corporation), and (2) the "earned income credit."

The last-named credit is allowed for the purpose of granting differential treatment in favor of earned income as against income from investments. From that part of the taxpayer's income which was earned there must be deducted any expenses incurred in earning such income. Certain arbitrary limits are then applied: if the taxpayer's net income is not more than $3,000, the entire net income is assumed to be earned, regardless of the source in the particular case, and if the net income is more than $3,000, at least $3,000 is assumed to be earned; earned net income in excess of $3,000 must be proved, and the total earned net income of any taxpayer is limited to $14,000, regardless of the source. These provisions obviously can result in only a very rough approximation to the actual earned net income in any given case. Ten per cent of the earned net income so determined is the "earned income credit." The remainder obtained by deducting these two credits (so far as applicable) from the "surtax net income" is the "balance subject to normal tax."

The rate of the normal tax is 4 per cent. As to the surtax, the first $4,000 of "surtax net income" is exempt; thereafter the rates are progressive from 4 per cent on the next $2,000 to 75 per cent on the excess of any income above $5,000,000.

An example. The important features of the individual income tax and the technical terms employed may be made clearer by a simple practical illustration. Let us take the case of a lawyer whose total receipts during the year 1938 consisted of: professional fees, $20,000, interest on bonds of the city of New York, $1,000, dividends from General Motors Corporation stock, $4,000, and a bequest of $10,000. He paid out during the year for professional expenses, $5,000, taxes, $1,000, and contributions, $1,000. He was married and had 4 children aged respectively 7, 10, 15, and 19 years.

It will be convenient to calculate first the "earned income credit." The net earned income in this case is evidently the $20,000 of fees less the professional expenses of $5,000, or $15,000. As this exceeds the maximum allowed as basis for the credit, we must take the statutory maximum, $14,000, as the net earned income. The credit is 10 per cent of this, or $1,400.

The tax for the year 1938 may now be determined by the following computation:

Professional fees,				$20,000	
Professional expenses,				5,000	
				15,000	
Dividends from General Motors Co.				4,000	
Total *gross income*,					$19,000
Taxes,				1,000	
Contributions,				1,000	
Total *deductions*,					2,000
Net income,					17,000
Personal exemption,				2,500	
Credit for 3 children,				1,200	
Total,					3,700
Balance (*surtax net income*),					13,300
Earned income credit,					1,400
Balance subject to normal tax,					11,900
Normal tax @ 4%,					476
Surtax	on first	$4,000		0	
	on next	2,000	@ 4%	80	
	on next	2,000	@ 5%	100	
	on next	2,000	@ 6%	120	
	on next	2,000	@ 7%	140	
	on last	1,300	@ 8%	104	
Total *surtax*	on	13,300			544
Total tax,					1,020

Incomes of corporations and partnerships. Corporations also (with certain exceptions) are taxed upon their net income. A corporation whose net income is not more than $25,000 pays a tax at rates ranging from 12½ per cent to 16 per cent according to the amount of the net income. A corporation with net income in excess of $25,000 pays a flat rate of 19 per cent with a credit of 2½ on the amount paid out in dividends. Corporations are subject also to a tax on capital stock and a tax on their "excess profits."

Partnerships as such are not taxable on their income. Their net income, whether actually distributed or not, must be allocated to the several partners and included by each in his individual income.

The reader will of course understand that the foregoing is only the briefest sketch of the most essential features of the United States income tax. The law itself is a lengthy document, loaded with technical details and administrative provisions of the utmost complexity. We have now before us however sufficient knowledge of the law to enable us to proceed to the discussion of the more fundamental principles involved, in connection with which we may draw some comparisons with the income tax systems of other countries.

Important principles exemplified in the United States income tax: Individual income tax base. The result of the personal exemptions and credits for dependents permitted in the United States is to give the individual income tax a very narrow base. On the incomes received in the year 1936, returns were filed (to August 31, 1937) by 5,334,513 individuals, of whom only 2,816,657 had taxable incomes. With a population of 128½ millions in 1936, returns were filed by slightly over 4 per cent of the population and income tax paid by less than 2¼ per cent. This does not mean of course that the income tax reached less than 2¼ per cent of the people in 1936. But making allowance for married couples and minors and other dependents, it is clear that the federal income tax reaches only a small minority — probably not over 10 per cent — of all the people.

In most other countries the personal exemptions and credits are much less generous than in America. Yet in any case, so long as any substantial minimum is exempted, the individual income tax is bound to rest upon a narrow base.

Tax-exempt securities. In the United States income tax, interest on state and municipal bonds and notes and interest on certain federal obligations are not taxable. There is a large and rapidly growing volume of these securities, furnishing a constant invitation to invest and escape taxation. Because of the high and steeply progressive rates of the federal income tax, the seductiveness of this invitation increases in direct proportion to the individual taxpayer's wealth.

Were the income tax imposed at a uniform rate (without progression) there would be some logic in exempting interest on the government's own obligations, since the government could presumably borrow at a correspondingly lower interest rate. With a progressive income tax there is no such simple relation between tax rate and interest rate. The government's gain in interest rates is determined by the value placed on exemption by those in the lowest income group which must be appealed to in order to market the entire debt. To them the tax exempt feature is not of great value. To the wealthy the corresponding escape from taxation is important. The government thus loses far more from tax exemption than it gains from lower interest payments.

There is of course no corresponding gain to the United States government to offset the exemption of interest on state and municipal obligations. This exemption is granted because of the belief that the taxation of such interest by the national government is unconstitutional. The federal income tax would be more just and the government's revenue more secure if the exemption of interest were abolished. This problem involves also the reciprocal exemption from state income taxes of the interest on federal obligations.

Relation between corporation and individual taxes. The taxation of the entire net income of corporations may be regarded as a form of double taxation to the extent of the part that is distributed in dividends to stockholders, since the latter are then taxable on their dividends as individual income. Until 1936 this was recognized, and a partial correction was made by requiring individuals to pay only the surtax (not the normal tax) on their dividend incomes. The effect of the present law is to place a specially heavy tax on income from investment in the stock of corporations.

The United States revenue act of 1936 introduced a novel feature in a special heavy tax, at rates from 7 to 27 per cent, on the undistributed earnings of corporations, in addition to the normal tax on all their taxable net income. The chief purpose was to force a greater distribution of earnings, retarding the accumulation of corporation surpluses, and bringing more income under the individual income tax. From the taxation standpoint, there is no justification for this device. As regards the corporations, the effect is to introduce an

influence upon the disposition of corporate earnings which may be contrary to the judgment of the responsible directors and managers. Special injury is likely to be done to corporations which by law or contract are restricted in their payment of dividends, to those that are seeking to retire debts or to pay off past losses out of current profits, and to the young, small, and growing businesses generally. The revenue act of 1938 went a long way toward removing this feature, by withdrawing the special undistributed earnings tax entirely from corporations with net income not over $25,000, while for the larger corporations it was reduced to a differential of only 2½ per cent. This vestige finally disappeared in the 1939 act.

The simple relationship between corporation and individual income taxation, conforming to the principles of justice and avoiding unnecessary interference with business management, would be a system in which corporations were taxed only on that part of their net income which was not distributed, at a flat rate reasonably corresponding to the rate paid by individual stockholders. Individual incomes from dividends would be treated the same as any other form of taxable income.

Progressive taxation of corporations. The revenue acts of 1935 and 1936 introduced another novelty in the form of progressive tax rates upon the incomes of corporations. Progression is generally applied in the taxation of individual incomes, upon the principle of distributing the burden in accordance with ability to pay. But this principle appears to have no valid application to the taxation of corporations. The equity of taxation imposed upon corporations can be judged only in relation to its effects upon natural persons. The burden of a corporation income tax rests chiefly upon the stockholders. If it were true that the large corporations were owned by wealthy stockholders and that the stockholders of the small corporations were poor or moderately well-to-do, there might be a rough justification for a progressive corporation income tax. But this premise does not agree with the facts. The shares of some of the largest corporations, such as the American Telephone and Telegraph Company, the United States Steel Corporation, and many others, are very widely distributed among small stockholders. On the other hand, many wealthy men are sole or chief owners of

incorporated businesses which, while representing substantial individual fortunes, are comparatively small as corporations go. Under such circumstances the graduated tax operates to defeat the ends of justice.

The graduated income tax doubtless receives some support from those who wish to discourage the large corporations. The rather questionable merits of this method of regulation would scarcely seem to justify the resulting injustice to the smaller investors.

Income tax rates. The World War brought a revolutionary increase in income tax rates. In Great Britain, for example, during the fifteen years before the war, the basic income tax rates varied between $3\frac{1}{2}$ and $6\frac{1}{4}$ per cent, with a super tax of $2\frac{1}{2}$ per cent after 1910. The present United States income tax was inaugurated in 1913 with progressive rates from 1 to 7 per cent. Today income tax rates have everywhere soared to hitherto unimagined heights. In the United States the present maximum rate (normal tax and surtax) of 79 per cent is the highest tax rate known to modern history, and to this must be added additional income taxes in about half of the states. It is to be noted that the act of 1935, which brought the surtax rates to their present climax, was actuated chiefly by political motives, expressing not so much the search for additional revenue as the administration's desire to support a "soak-the-rich" program. Indeed the social motive of equalizing fortunes is never entirely absent from the income tax and shares with the theory of taxation according to ability responsibility for the highly progressive rates of modern income taxes.

Since the War the practicability of such rates has come under serious question, especially in the United States. It is possible for income tax rates to be so high as to discourage investment in industry. Investors and entrepreneurs may conclude that they are not warranted in taking all the normal risks of industrial enterprise and in addition facing the prospect that in case of successful ventures the government will take the major part of the profits, while the burden of losses will fall entirely upon them. During the past few years this has been the attitude of many American investors and business men, and it is becoming clear that this has been one of the major causes of the continuation of industrial depression. In

consequence the tax revenue is less than it probably would be with lower rates; in other words, the rates have been pushed beyond "what the traffic will bear."

There is another important aspect of excessive income tax rates. Even with the best possible administration, which the United States confessedly does not have, the success of an income tax must depend materially upon a prevailing spirit of fairness and coöperation on the part of the taxpayers. So long as this spirit is general, administrators of the law and the courts can cope with the actual criminal evaders. When this spirit disappears, when there is a general feeling that only the quixotic few pay their full taxes, when the majority become lawbreakers, then no administrative machinery that human intelligence can devise will be able to master the problem. This condition is threatened if rates are made too high.

Elasticity of the income tax. In the early days of the income tax, it was rightly regarded as a peculiarly elastic source of revenue; its yield could be varied at the will of the taxing authority by raising or lowering the rates. For more than half a century the income tax thus rendered conspicuous service to Great Britain. It permitted easy adjustment of the budget to the normal fluctuations in government needs and in the product of other less elastic sources of revenue; and it came strongly to the nation's support in emergency, as on the occasion of the South African War. During the World War, the income tax was called upon to the limit, especially by Great Britain and the United States.

But elasticity must be permitted to work both ways; otherwise its loss is threatened. Income tax rates were somewhat reduced after the World War but still remained far above their pre-war level and close to the limit of elasticity. This condition was rudely disclosed with the coming of economic depression after 1929. The financial exigencies of government led to increased demands upon the income tax, to which the tax was not able to respond. The income tax is peculiarly sensitive to general economic conditions. If in time of depression this tax is to be called upon for an increased contribution, then rates must be raised drastically in order to counteract the decline of the tax base. But with rates already close to the limit of elasticity, sufficient increase was not possible. In

the United States, for example, even with increased rates, the yield of the income tax fell off disastrously. In 1930 the income tax produced 2,410 million dollars, contributing 58 per cent of the total revenues of the federal government. In 1933 its yield was 746 millions, being only 36 per cent of a total revenue which had itself dropped to less than half the level of 1929. In spite of every effort of the government, the yield of the income tax fell off by more than two-thirds just when it was needed most, and its failure was chiefly responsible for the collapse of the federal revenue system in this emergency.

Shifting and incidence of taxes on property and income: The tax on economic rent or land value. We have already learned that shifting of the burden of a tax by the immediate payer to someone else is conditioned upon a power, given by the tax itself, to raise the price of some commodity or service which the immediate payer has to sell. The simplest theoretical example of this principle is that of the tax upon the landowner's income from economic rent. Having in mind the nature of economic rent (which we have investigated in Chapter XVI) we readily reach the conclusion that a tax upon economic rent gives the landowner no power to increase any price he is charging others. If the land is rented, the tenant (assuming he is already paying the full economic rent) can pay no more rent without being forced to operate at a loss Rather than do this he will leave this land and either go out of business or betake himself to the marginal land, where there is no rent and no tax on rent. Neither the tenant nor the landowner who is using his own land is given any power to charge a higher price for his crop by a tax on economic rent. The price is determined by demand and the cost of production on the no-rent land, neither of which is affected by the tax in question. We conclude that a tax on the economic rent of land is not shifted but is borne by the immediate payer, the landowner.

One necessarily reaches the same conclusion as to a property tax upon land alone, since the value of the land is the capitalized value of its economic rent. A property tax on land can certainly give the owner no power to sell the land at a higher price; nor does it give him power either to exact more rent from a tenant or charge

higher prices for agricultural products, since the marginal land, having no economic rent and so no value, is no more affected by a property tax than by a tax on rent.

Capitalization of taxes. The tax on economic rent or on land value furnishes the simplest illustration of the principle of "tax capitalization." If an acre of land is yielding an economic rent tax-free of $10 a year and the rate of interest at which such incomes are discounted is 5 per cent, the land will be worth $200. If now, assuming no other material change in the tax system, an annual tax is imposed on economic rent at the rate of 10 per cent, the landowner will have left an income of only $9 a year, which, capitalized at 5 per cent, makes his acre of land worth $180. Taking 10 per cent of the annual income is equivalent to taking in one lump sum 10 per cent of the value of the land. No one will buy this land for more than $180, and the purchaser, receiving $10 in rent and paying $1 in tax, obtains 5 per cent on the investment, the same as he would have received if he had paid $200 for an acre not subject to the tax on its rent. Of course precisely the same result follows the imposition of an annual property tax on the value of the land. Taxes on economic rent or land value thus tend to rest as a permanent burden upon those who owned the land at the time of the imposition of the taxes and not to be passed on to later purchasers. The principle of capitalization applies, not only to land, but to any other capital instrument where it is not possible to shift the tax; for example a tax on bonds.

Taxes on dwelling houses or their rent. The rent paid for a dwelling house is distinguished from the rent paid for the location by the obvious fact that buildings are constantly wearing out and being replaced with new ones, whereas land is permanent and practically fixed as to quantity. Before the imposition of a tax on the rent of dwelling houses, there was a certain normal return obtainable from renting dwelling houses; now that return has been reduced by the amount of the tax. Investors will tend to stop building new houses, preferring to put their capital into other ventures where the normal return is obtainable. The diminished supply of houses for rent, as old houses wear out and are not replaced by new ones, will cause an increase in the rents which house

owners can exact and tenants will pay, until the rents are high enough to encourage new building. Our general conclusion then is that the final incidence of a tax either on dwelling houses or on their rent tends to be upon the occupiers, whether they be owners or tenants, the tax being shifted when imposed upon the owners of rented houses.

The tax on business property. The property tax on wealth used in business, such as factory and store buildings, warehouses, railway equipment, machinery, goods and materials, etc., generally means an addition to the cost of the products of business enterprise. Under competitive conditions anything that increases cost of production will tend to raise prices Even under monopoly conditions, there will generally be some tendency to raise prices. It is thus that a property tax upon wealth, other than land, used in production is generally shifted through higher prices of products and so finally rests as a burden upon the consumers of the products.

The property tax on consumers' goods. When the property tax falls upon goods, such as household furniture, books, jewelry, automobiles, etc., in the hands of the final consumers, there is no longer any question of a price to be charged and hence no possibility of shifting the tax.

The income tax. In the case of the tax on competitive business profits there may be a certain degree of shifting through the charging of higher prices. But since the tax encroaches only upon net income or profit, and takes only a fractional part of that, an income tax at moderate rates will not generally drive business men to sacrifice such profits as remain to them. Very heavy rates, such as have recently been imposed in the United States, undoubtedly do deter investors and entrepreneurs from risking their capital in business and so tend to raise the prices of the products of industry.

In the case of a monopoly an income tax upon profits cannot possibly be shifted. The monopoly, having already set its price so as to reap the maximum profit, has nothing to gain by a change in price, even though it face a tax up to 100 per cent of its profits. It is better to retain any given fraction of a larger net income than the same part of a smaller income which would result from a change in price.

Upon non-business incomes the possibility of shifting is even less than in the case of business incomes. Ordinarily there is here no power to increase any price on account of the tax and so no way of shifting its burden.

The poll tax. This is a tax, generally at a uniform rate of one to three dollars, upon every adult person, or every voter, or some other broadly inclusive class of the people. It is expensive to collect, except from those who are already subject to the property or income tax, and while it is (on paper at least) a recognized part of the American local tax system, it is generally honored more in the breach than in the observance, and the revenue derived from it is usually small.

Of late the poll tax has come into somewhat more favorable recognition from two viewpoints. Appreciation of the great importance of having everyone feel some financial stake in his government and some direct interest in its finances has led to the proposal of a universal income tax, which would allow no personal exemption and would start with a minimum tax of from one to five dollars. Such an income tax would obviously absorb the existing poll tax. This plan has much to commend it. Thus far it has produced only a few tentative steps in actual legislation.

On the practical side, the poll tax has recently proved itself useful in bringing home to the people financial responsibility for certain government expenditures conferring special benefits upon specified classes of the people. Thus the State of Massachusetts, after the World War, granted a bonus to war veterans and raised the necessary funds by means of a state poll tax. Connecticut in 1935 inaugurated a system of old age pensions, with provision for raising the money by means of a town administered poll tax.

Consumption taxes. Next in importance to the property tax and the income tax comes a group of taxes whose central characteristic is the exaction of a tax contribution from the consumers of certain goods in proportion to their consumption. They appear in a great variety of forms, sales taxes, excise taxes, licenses, customs tariffs. For the purpose of our study we may divide consumption taxes into two main groups: (1) excise taxes (including sales taxes and business license taxes), and (2) customs tariffs.

Excise taxes: General account. An excise is in general a tax levied upon the manufacture or sale or consumption of commodities or the exercise of an occupation or enjoyment of a privilege within the limits of the jurisdiction imposing the tax. What will chiefly interest us are the taxes levied upon the domestic manufacture and sale of commodities, such as are commonly known in the United States as "internal revenue" taxes when imposed by the federal government and "sales taxes" when imposed by the states.

Excise or sales taxes played only an insignificant rôle in the United States until the time of the Civil War. Starting in 1860, Congress had before the end of the war built up an elaborate system of sales taxes, covering nearly every commodity of wide consumption, levied now upon manufacturers, now upon dealers, often involving double or multiple taxation of the final product, and carrying a wide variety of rates.

After the war this elaborate system was generally repealed, leaving as a permanent part of the federal revenue system only the excises on liquors and tobacco and a few minor items, such as playing cards, oleomargarine, etc., whose retention was largely for sumptuary purposes or to favor the producers of a competing commodity. The liquor and tobacco taxes are levied upon the manufacturers. At the time of the Spanish War in 1898–99 and again during the World War, the United States government obtained large additional revenue from excise taxes upon other selected commodities and services. On each occasion these excise taxes were generally repealed soon after the war. The depression following 1930 similarly occasioned some fresh imposition of excise taxes.

Excise taxes are generally used by national rather than by local governments. This is principally because of the tendency of such taxes to drive business outside the borders of the taxing jurisdiction. In the United States the several states have generally been deterred from extensive use of excise taxes by these considerations, coupled with the constitutional prohibition of customs tariffs and excises burdening interstate commerce. The states are not prevented however from levying moderate excise taxes upon industries

catering to a strictly local market or upon local or retail sales, though caution has always to be exercised lest the people be driven to make their purchases outside the state.

One of the most striking features of post-war finance all over the world is the rapid and widespread development of sales taxation, forced by the fiscal needs of governments harassed by increased demands for expenditure and dwindling sources of revenue. Among the American states there is to be witnessed a really extraordinary development of this form of taxation. The excise tax on gasoline is now universal. Excise taxes or license taxes are imposed upon sales of liquor by all states which do not prohibit its sale. As to consumption generally, by the close of 1937 more or less general sales taxes were in force in some twenty-five states.

Shifting and incidence of excise taxes. The excises illustrate the operation of shifting in its simplest form. Consider for example the case of the United States internal revenue tax upon cigars. The tax is imposed upon all manufacturers of cigars at a certain rate per hundred and is thus an addition to the normal unit cost of production. Manufacturers are therefore able to increase the selling price by practically the amount of the tax and so to pass the burden on to the wholesale dealers to whom they sell their product. These in turn pass the tax on to the retailers, who likewise pass it on in the prices charged the consumers. Sometimes an excise tax may be imposed upon the dealers, wholesale or retail, instead of upon the manufacturers, or upon the consumers themselves at the time of purchase of the taxed articles, the law requiring the dealer to collect the tax and pass it back to the government. In all such cases the price paid by the consumer is of course increased by the amount of the tax. When the excise is based upon an industrial occupation, instead of upon a commodity, the same result follows; the tax becomes an added cost of conducting the particular occupation and is passed on to the consumers by means of a higher price for goods or services. The retail liquor seller's license is a familiar example.

Effect on production. It is not to be inferred that an excise tax is wholly without burden upon the manufacturers and dealers; indeed the reverse is generally true, due to the normal shape of the

curves of demand and supply. Let *dd* and *ss* (Figure 1) represent respectively the demand and supply curves of a certain commodity, determining the selling price *pa*. Now suppose an excise tax is imposed upon the manufacturers or dealers. The unit cost being thereby increased, there will be a new supply curve *s′s′*, higher than the former curve. The intersection of this curve with the demand curve, of course not changed by the tax, will determine a new price *p′b*. This price is higher than the old price, but except in the case of constant cost industries it is generally not increased by the full amount of the tax. What is of more importance, the quantity sold has fallen off (from *Oa* to *Ob*). The manufacturers and dealers will,

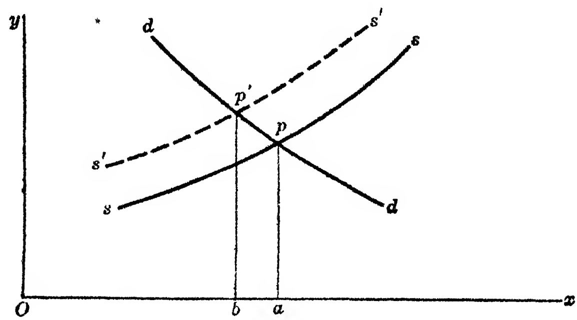

FIG. 1

after the process of shifting, be able to charge enough to give them the ordinary rate of return, but only on a smaller volume of business.

Business taxes. Excise taxes imposed with respect to the conduct of business are not uncommon, though less important than those forms of taxation which have thus far been studied. License taxes upon individuals and corporations engaging in certain specified occupations are common in the Southern states of the United States and are often quite far-reaching and productive of considerable revenue. Such taxes are not important in the more advanced industrial states, and their significance is likely to decline as time goes on. State and local license taxes upon retail liquor dealers are a well-known example of the business tax. On the whole the special field of business taxation is not very deeply cultivated in the United States, though there are some indications of an increasing interest.

The customs tariff. Customs or tariff duties are taxes imposed upon those who bring goods across the frontier, the base being the goods so moved. The Constitution of the United States prohibits all duties on exports and also practically forbids the levy of any customs duties by the states. From the beginning of the American nation in 1789 to the Civil War, the tariff was almost the only tax imposed by the national government, and from then to 1913 it divided honors about equally with the excise or internal revenue taxes. Today the tariff yields about one-tenth of the federal revenue.

Economic effect of the tariff. It may help the reader to grasp the salient characteristics of the tariff, together with their immediate consequences, if we set up a simple numerical example. Let us suppose that a certain kind of cloth is being imported, free of duty, and sold in the home market at $1.00 a yard and that 1,000,000 yards are sold in a given year. We assume that the cost of making this cloth at home is $1.50 a yard. Obviously the reason for the import is that this is the cheapest way to obtain the cloth, and of course there will be no home manufacture. The simple situation resulting is shown in column 1 of the following table:

	1 No Tariff	*2* 10% Tariff	*3* 50% Tariff	*4* 75% Tariff
Number of yards imported	1,000,000	900,000	300,000	. .
Number of yards domestic	.		300,000	600,000
Total number of yards consumed	1,000,000	900,000	600,000	600,000
Price per yard to consumers	$1.00	$1.10	$1.50	$1.50
Cost to consumers	$1,000,000	$990,000	$900,000	$900,000
Additional cost because of tariff	.	90,000	300,000	300,000
Revenue to government	.	90,000	150,000	. . . .
Additional cost in excess of revenue	. . .	. . . . .	150,000	300,000

Now suppose that next year a tariff is in effect at the rate of 10 per cent, which will make the cloth cost $1.10 in the home market. This price being still less than the home cost, there will be no home manufacture, and all the cloth consumed will still be imported. In accordance with the normal demand curve, and assuming other circumstances unchanged, we assume that the amount consumed will fall off somewhat with the increase of price and that 900,000 yards will be taken in this year. The result is shown in column 2

of the table above. Consumers pay $990,000 for cloth which, were it not for the tariff, they could have obtained for $900,000, and this additional cost goes to the government in revenue. Here is obviously the case of a simple tax on imports.

Next year a 50 per cent tariff is in effect, and the imported cloth will be sold, if at all, at a price of $1.50. This price is just equal to the cost of making the cloth at home, and the foreign and domestic products are in competition on an equal basis. Presumably they will divide the market between them in some ratio, which for the sake of simplicity we may assume will be an equal division. The result is shown in column 3 of the table, assuming a still further drop in the amount taken on account of the rise in price to $1.50. Consumers now take 600,000 yards, paying for them $300,000 more than they would have paid for this quantity had there been no tariff. This additional cost however does not now all go to the government; the government collects 50 per cent upon the 300,000 yards imported, giving it a revenue of $150,000; but it gets no revenue from the 300,000 yards of domestic manufacture. We meet here for the first time the phenomenon of *protection*, by which we mean *encouragement given to a home industry by means of a tariff*.

Now let us proceed to the last case, that of the *prohibitive* tariff. A duty of 75 per cent being in effect, the foreign cloth can be sold only at a price of $1.75. Since the home industry can furnish the cloth (at least 600,000 yards according to our assumption) for $1.50, there will be no possibility of selling the imported article. This situation is shown in column 4 of the table So far as the consumers are concerned, the situation is the same as in the previous case, except that all their purchases are of the domestic cloth. The government however now gets no share at all of the $300,000 additional cost imposed by the tariff upon the consumers. This is the extreme of protection.

To the reader who has fully grasped the meaning of the foregoing example, the most important principles of the customs tariff will now appear almost as a matter of course. They may be summarized in the following language (the reader will find it helpful to refer again to the numerical example above):

(1) A tariff will cause the price of the dutiable article to rise in the home market by about the full amount of the duty, up to a duty which is equal to the difference between the home and the foreign unit cost of production.[1] Above that the duty is prohibitive, and the further result depends on whether there is competition or monopoly in the home industry. If competition prevails, further increments of duty have no effect on the price; under monopoly conditions the price may rise by more than the difference in unit production costs up to the full amount of the duty.

(2) A tariff imposed at a rate less than the difference between the home and the foreign unit cost of production will have no immediate effect on home industry and will produce a revenue equal (disregarding costs of administration, etc.) to the added burden placed upon consumers.

(3) A tariff equal to or greater than the difference in unit production costs gives protection, and the additional burden placed upon consumers will be in excess of the revenue obtained by the government. If the tariff is in excess of the difference in unit production costs it is prohibitive, in which case the additional burden on consumers is accompanied by no revenue whatever to the government. Stated in slightly different words, any tariff which gives protection to domestic industry necessarily causes a loss to consumers greater than the gain obtained by the government.

(4) The additional burden placed upon consumers by a tariff which gives protection will be equal to the difference between the home and the foreign cost of producing all the goods (domestic and foreign) consumed, except when a prohibitive tariff accompanies a monopolistic condition in the home industry. In that case, the consumers may be burdened by an amount in excess of the difference in the entire cost of production.

The revenue tariff. The traditional country of "free trade" was Great Britain. From the reform of her tariff system in the middle of the nineteenth century until after the World War, Great Britain

[1] The word "about" is introduced in this sentence in order to allow for the possibility of a change in the foreign cost of production resulting from the decreased purchase of the imported goods. The practical importance of this would generally be slight.

followed a policy of tariff for revenue with virtually no protection. In order to make her tariff truly a revenue producer without disturbance from protective influence, Great Britain adopted certain devices which are universally effective for this purpose. In the first place, her duties were placed upon a comparatively small number of articles, these being generally commodities which in the nature of the case could not profitably be produced in Great Britain even in the face of high tariff duties. Such are tea, coffee, sugar, cocoa, spices, various fruits, and other tropical products. The second device for safeguarding the revenue tariff from defeat through the checking of imports and the development of home industries is to impose an internal excise duty on each home-made good to match the tariff duty on the imported article. Finally there is the extreme safeguard of prohibiting outright the home industry, which is employed by Great Britain in the case of tobacco growing. By such means as these it is possible to build up a tariff system of extreme simplicity, easy to administer, having no disturbing influence upon the country's trade and industry, taking nothing from the people beyond what goes into the public treasury, and withal yielding a handsome revenue.

Fiscal weakness of the protective tariff. Few countries however outside of Great Britain have chosen this kind of tariff. Virtually everywhere protection is consciously sought to a greater or less degree, and the employment of devices to neutralize the protective tendency of duties imposed for the sake of revenue is rare. At the extreme protective pole stands the United States, where, in spite of the great importance of the tariff as a source of revenue, the protective purpose has always been strong or even predominant.

Throughout the history of the United States, the tariff revenue has been notoriously irregular, out of relation to governmental needs, and not amenable to legislative control. In 1807 the tariff revenue was double the year's expenditures; in 1837 the tariff, still the sole tax revenue of the federal government, produced less than one-third of the amount necessary to cover the year's expenditures. The tariff failed dismally to meet the war emergencies of 1812 and 1861. Since the Civil War the steadying influence of the excise taxes has somewhat mitigated the gyrations of the tariff revenue,

but throughout the entire national history up to the World War dependence upon the tariff has been, more than anything else, the cause of that succession of uncomfortable deficits and embarrassing surpluses which testify to an imperfect financial system.

The explanation is simple. The tariff being looked upon as an instrument for giving protection to certain industries, its rates cannot be raised or lowered to meet fiscal conditions without opening up the whole political conflict over protection. Furthermore the relation between tariff rates and tariff revenue is extremely complex, and it is impossible to predict accurately the effect of a change. By checking the flow of imports an increase of rates may diminish revenue instead of increasing it; a reduction of rates may increase revenue.

Death taxes: General description. The death of the owner gives to government a convenient opportunity to impose taxes from time to time upon the transfer of property of every description, an opportunity which is availed of by sovereign governments quite generally the world over.

The typical modern *inheritance tax* contains the following technical features. At the outset property left to certain religious, educational, or charitable institutions is generally entirely exempt from the tax. Of the remaining taxable estate, a certain amount is usually exempt, this amount varying according to the degree of relationship between the decedent and the beneficiaries. Above the exempt minimum, the tax is calculated at rates which are usually doubly progressive: (1) the rates start low for bequests to the closest relations and increase step by step as the degree of relationship becomes more remote, and (2) within each class of beneficiaries, the rate is progressive according to the amount transferred. Among the states of the United States the top rate on direct heirs varies all the way from 3 per cent to 16 per cent; on the most remote relations and strangers the maximum rates run as high as 40 per cent.

A tax levied upon the entire estate without relation to the amounts of the distributive shares or the relationship of the beneficiaries is called an *estate tax.* As an example we may take the United States estate tax, as amended by the revenue act of 1935.

The tax, as regards resident decedents, is imposed only upon estates whose amount exceeds $40,000 and is based upon the "net amount" of the entire estate, calculated by subtracting a lump sum of $40,000 besides other deductions. With respect to the net estate so defined the rate begins at 2 per cent upon the first $10,000 and increases progressively until the maximum rate of 70 per cent is reached, applying to that part of the net estate in excess of $50,000,000.

Some fundamental questions. A bequest, except where received by a dependent of the testator, comes as a more or less extraordinary windfall, different from one's ordinary regularly recurring income. This is usually taken to justify quite heavy taxes upon estates and inheritances as compared with the other kinds of taxes on property, income, etc. Similar conclusion is drawn from the prevailing concept of property rights, which holds them to be distinctly less sacred when it comes to disposition of one's possessions after death, especially that part which is not passed on to those immediately dependent upon the testator. There are those who would go still further and use the inheritance tax as an instrument to check the rise or the perpetuation of large family fortunes, and there can be no doubt that the very high rates in the upper brackets of some death taxes — witness the United States estate tax, and especially its amendment in the revenue act of 1935 — are partly in response to a demand for equalizing wealth by heavy levies on the large fortunes.

Death tax rates, at least as regards bequests to indirect heirs and strangers, might be pushed very high before reaching any limit set by the principle of ability to pay. But before this limit is reached there is almost certain to be a limit set either by administrative considerations or by broad social policy. There are many kinds of business arrangements and legal devices whereby death taxes may be escaped while still keeping within the letter of the law. There is always the possibility of escape through gifts before death, a door which no development of gift taxes is likely completely to close. The higher the rates the more will such avenues of escape be resorted to. The second limit is reached when the tax becomes so high as to discourage the accumulation and husbanding of capital

or to interfere with its most productive use. While it is probable that this effect of the inheritance tax has been frequently exaggerated, there would eventually come a time when the public loss from decline of capital and disturbance to business and investment would exceed the possible benefits from the revenue thus contributed.

XXXII

PUBLIC FINANCE

Cost of government and national income. In the year 1938 there were expended by government in the United States about 17 billions of dollars. The federal government's share was 7⅝ billions; the 48 states spent about 4⅓ billions; and the various local governments in all the states accounted for about 5½ billions.[1] These are huge sums, beyond the power of the imagination to visualize. Somewhat easier to grasp is the fact that government expenditure in that year took more than a quarter of the total income of the people of the United States, estimated at 62¼ billion dollars. If perchance the situation in 1938 may be considered abnormal, on account of depression and "New Deal" expenditures, we may go back to the year 1929, when the total expenditures of government were about 12 billions and represented about 15 per cent of a national income of 79½ billions. It will be noted that between these dates there have been both a drastic decline in the people's income and a heavy increase in the cost of government.

Growth of public expenditure: General summary. The constant growth of public expenditures is an outstanding phenomenon of modern history. Starting as far back as reliable records will take us, this universal tendency appears. It applies to all types and grades of government: monarchies and republics, peaceful nations like Switzerland as well as militaristic Germany, imperial governments, national governments, states, and municipalities. The increase in

[1] These figures are not exact In particular the state and local expenditures have to be estimated The latest official estimates are those of the United States Treasury Department, published in *Bulletin of the Treasury Department*, August, 1939, and summarized on page 4 of this bulletin. These estimates are as follows:

Expenditures of the federal government	$ 7,626,000,000
Expenditures of the state governments	4,223,000,000
Expenditures of the local governments	5,621,000,000
Expenditures of all grades of government	$17,470,000,000

These figures and corresponding figures in various other tables in this chapter are not precisely comparable, owing to certain differences in the method of calculation, which however are not great enough to affect any of our conclusions

government expenditures before the nineteenth century is on the whole more apparent than real, the great advances in the absolute figures being substantially offset by such factors as diminishing reliance on goods and services directly rendered, growth of population, increase in territory, increased wealth of the people, and depreciation of the monetary standard. But since the beginning of the nineteenth century there has undoubtedly been real growth.

United States federal government. For the greater part of the nineteenth century the paucity and unreliability of the statistics of American state and local expenditure restrict us to the study of the expenditures of the federal government alone.

ORDINARY EXPENDITURES OF THE UNITED STATES GOVERNMENT [1]

Year	*Total expenditures*	*Per capita expenditures*
1792	$ 5,079,532	$ 1.28
1800	10,786,075	2.03
1810	8,156,510	1.13
1820	18,260,627	1 89
1830	15,143,066	1 18
1840	24,317,579	1 42
1850	39,543,492	1 71
1860	63,130,598	2 01
1870	309,653,561	8 03
1880	267,642,958	5 34
1890	318,040,711	5 05
1900	520,860,847	6 85
1910	693,617,065	7 54
1915	760,586,802	7 66
1919	18,514,879,955	176 33
1920	6,403,343,841	60.57
1925	3,063,105,332	26 60
1929	3,298,859,486	27.15
1930	3,140,268,884	27 93
1931	3,651,515,712	30 47
1932	4,711,015,145	38 95
1933	4,681,348,827	38.55
1934	6,745,185,992	54 45
1935	6,802,266,916	51 42
1936	8,476,558,108	66 00
1937	8,001,187,347	61 90
1938	7,700,909,327	59 20
1939	9,210,091,581	70.31

[1] The figures for total expenditures are from *Annual Report of the Secretary of the Treasury on the State of the Finances*, 1937, pp. 350–353, and (for 1938 and 1939) *Daily Treasury Statement*. Postal expenditures are included only to the extent that they exceed the postal revenues.

For the first seventy years of the national existence of the United States, the expenditures of its government were on a modest scale, the increase from eleven millions in 1800 to 63 millions in 1860 being no more than sufficient to keep pace with the growth of the country's population, as is shown by the figures for per capita expenditure. The figures for 1870 show the effect of the Civil War cost, and the later figures give evidence of the universal truth that war not only adds a war-time burden of extraordinary costs but sets a new scale of peace-time expenditure from which no people has ever been able to return to the pre-war status. Other influences contributed further to the growth of public expenditures, so that from the Civil War to the World War the cost of the federal government averaged more than four times as much as before the Civil War, after making full allowance for the growth of population.

The gigantic expenditures of the World War reached their climax in 1919. Expenditures then dropped, but they were never again as low as three billions. The new level reached in 1925 was four times as high as the expenditures of 1915, or three and a half times as high after allowing for the increase in population. In 1932 begins another decided upward swing, the result chiefly of depression and "New Deal" expenditures, which will be analyzed at a later point in this chapter.

All grades of government in the United States. For the more recent years we have data enabling us to bring state and local expenditures into the picture of the growth of government costs in the United States. The following table is the result of researches of the National Industrial Conference Board.[1]

TOTAL GOVERNMENTAL EXPENDITURES, UNITED STATES

Total (millions)

	1890	*1913*	*1929*	*1936*
Federal	$ 291	$ 692	$ 3,046	$ 8,547
State	77	383	1,943	2,433
Local	487	1,844	6,720	6,038
Total	$855	$2,919	$11,709	$17,018
Nat'l income	$10,701*	$31,450	$79,498	$65,246

[1] *Cost of Government in the United States, 1926–1927*, 1928, pp. 2, 63; *Cost of Government in the United States, 1933–1935*, 1936, pp. 4, 30, *Cost of Government in the United States, 1935–1937*, 1938, p 4, and *National Income in the United States, 1799–1938*, 1939, pp. 6–7.

TOTAL GOVERNMENTAL EXPENDITURES, UNITED STATES (*Continued*)

Per capita

	1890	*1913*	*1929*	*1936*
Federal	$4.61	$7.17	$25.06	$66.55
State	1.22	3.97	15.99	18.94
Local	7.72	19.10	55.30	47.01
Total	$13.56	$30.24	$96.35	$132.51
Nat'l income	$173 00*	$326 00	$654.00	$508 00

Ratio of national expenditures to national income

1890	*1913*	*1929*	*1936*
8.0%*	9 3%	14.7%	26.1%

* National income as of 1889.

For the year 1938 the National Industrial Conference Board offers the following preliminary estimate: federal, 7.5 billions; state and local, 9.3 billions; total, 16.8 billions.

It will be noted that division of the total cost of government between the three grades of government had not greatly changed in the generation before the depression. The federal government accounted for about one-third of the total, the states for ten to fifteen per cent, and the local governments for something more than half. This is significant as showing that, in spite of the powerful influence of war in causing expansion of federal expenses, the costs of state and local government had not fallen behind. The climax of total government cost was in 1919. After that the federal government reduced its expenses. But in the meantime the expenditures of the state and local governments continued to rise, and after 1925 those of the federal government were also growing, so that the total cost of government was again on the increase. The effect of the depression was to increase substantially the relative share of the federal government in the total and to reduce correspondingly the share of the local governments. In 1936 the distribution was: federal government, 50.3 per cent; states, 14.3 per cent; local governments, 35.4 per cent. According to preliminary estimates of the National Industrial Conference Board, the distribution in 1938 was: federal government, 45 per cent; state and local governments together, 55 per cent.

In 1890 a sum equal to 8 per cent of the income of the American

people was devoted to defraying the cost of government. In 1929 the percentage was about fifteen. In 1936 the effects of depression were apparent, drastically reducing the incomes of the people and increasing rather than decreasing the cost of government. The ratio had risen to about 26 per cent. The profound significance of this development should be apparent. As time goes on a continually greater part of the income of the people is being taken and employed by government. In return valuable services are being rendered by government. Whether this distribution of the national income is giving the maximum of utility to the people and, if so, how far the portion devoted to government may continue to increase without a decrease in utility are questions to which the student may well give serious thought. They raise problems in economics, in political science, and in social philosophy which, though most inviting, must be regarded as outside the field of this elementary treatise.

Functional analysis of government costs in the United States. In discussing the cost of government we shall have to direct our attention chiefly to the United States. As of all national governments, the largest item of expenditure is for defense — war. Besides war the leading expenditures of the United States government are rivers and harbors and other public works, marine and land transportation, various facilities for national industry, maintenance of justice and security, regulation of private business, sanitary regulations, immigration and naturalization service, and the maintenance of the legislative, judicial, executive, and administrative branches of the government. The table on the opposite page is a summary functional analysis of the expenditures of the federal government in the fiscal year 1937.[1]

This table, while representing fairly the character of the federal expenditures of recent years, is not typical of conditions before 1932, owing to the fact that since that year the expenditures of the United States government have included very large sums devoted to efforts to combat economic depression and upon the various projects of the "New Deal." These expenditures will be further considered later in this chapter.

[1] National Industrial Conference Board, *Cost of Government in the United States, 1935–1937*, 1938, pp. 17, 19.

EXPENDITURES OF UNITED STATES GOVERNMENT, 1937

Function	*Millions of dollars*	*Per cent of total*
Regular operating expenditures		
Legislative, judicial, and civil establishments	814.7	10.2
National defense	888.6	11.1
Veterans' pensions and benefits	1,128 2	14.1
Interest on the public debt	866 4	10.8
Other	48 1	0 6
Total regular operating expenditures	3,746.0	46.8
Public Works		
Public highways	350.6	4.4
Public buildings	76 3	1.0
Rivers and harbors, improvement	142 4	1.8
Grants to public bodies, including administration	272.9	3.4
Other	237 2	3.0
Total public works	1,079 4	13.5
Relief		
Direct relief	184 3	2.3
Work relief (WPA and CWA)	1,896 7	23 7
Civilian Conservation Corps	385.8	4 8
Total relief	2,466 8	30.8
Miscellaneous		
Loans (net excess of credits)	−307 1	−3 8
Subscriptions to stock and surplus	47 1	0 6
Agricultural adjustment program	515 8	6 4
Social security	447 7	5 6
Railroad retirement	5 5	0 1
Total Miscellaneous	709 0	8 9
Total expenditures	8 001 2	100 0

Education stands first among the expenditures of American state and local governments Next come highways, care of dependents and defectives, maintenance of justice and security, protection against disease, and interest on debt. The statement on the next page of expenditures of the states, the counties, and the other local governments combined was prepared by the United States Census.

Paying for government services. Obviously in order to meet its expenditures, every government must have corresponding receipts. In the two chapters immediately preceding this one, we have analyzed the various types of income which are available to government and have studied in some detail the more important forms

of taxation. In any particular government we find a *revenue system,* made up of a collection of taxes and other sources of income and presenting characteristic problems.

GOVERNMENT COST PAYMENTS, STATE AND LOCAL GOVERNMENTS, 1932[1]
(In thousands of dollars)

Function	*States*	*Counties*	*Other local divisions*	*Total*
General departments:				
General government	123,333	251,150		
Protection to person and property	86,875	44,231		
Health and sanitation	37,292	32,778		
Highways	235,690	236,350		
Charities, hospitals, and corrections	274,920	182,120		
Schools	602,065	178,406		
Libraries	2,405	3,772		
Recreation	8,054	7,618		
Development and conservation of natural resources	72,965			
Miscellaneous	57,794	44,767		
Total general departments	1,501,393	981,192	3,566,523	6,049,108
Public service enterprises	6,330	489	301,125	307,944
Interest	112,258	118,875	612,735	843,868
Outlays	885,854	311,270	1,164,571	2,361,695
Total	2,505,835	1,411,826	5,644,954	9,562,615

The United States federal revenue system. Light will be thrown upon this subject by a brief historical sketch of the revenue system of the United States federal government. The American colonists were always sensitive to taxation. Taxation was a sore point between them and Great Britain, and when the Revolution released them from British taxation they were loathe to give to their new national government adequate taxing powers. The Articles of Confederation broke down largely because they gave the government no power to tax the people, and the Constitution of the United States contained important restrictions upon the taxing power of the national government. Chief among them was the provision that direct taxation may be employed only if apportioned among the states according to their population. This rule made the practical task of levying a direct tax almost prohibitively difficult. Some timid experiments were made soon after the adoption of the Con-

[1] United States Bureau of the Census, *Financial Statistics of State and Local Governments 1932,* 1935, p. 28.

stitution and again at the time of the Civil War. But these direct taxes were ineffective and produced little revenue, and the attempt has never been repeated.

From the start the national government relied upon indirect taxation. And of the two principal types of indirect taxes, the customs tariff was at the first chosen rather than the more obvious internal excise taxes, again on account of popular antagonism to taxation. While there was some half-hearted use of excise taxes up to the close of the War of 1812, the tariff was practically the only tax resource of the national government until the time of the Civil War. It proved, as we have learned, an erratic source of revenue at best and failed completely to meet the emergency needs of war.

The Civil War was of such duration as to give the people time to overcome their traditional repugnance to internal taxation. Starting in 1862, excise taxes were enacted which finally laid burdens upon almost every article of industry and commerce, and in the financial reorganization after that war, the excise system, though greatly reduced by freeing the majority of taxed commodities, was not entirely abandoned as it had been after the War of 1812. A group of excises, principally upon liquors and tobacco, remained and provided not only an important source of peacetime revenue but also the skeleton upon which it was possible to build quickly an effective system of war taxes in time of future need, of which good use was made on the occasions of the Spanish War and the World War.

For the half century following the close of the Civil War, the tariff and the excises shared about equally (tariff revenue being generally a little greater than excise revenue) the task of producing the revenue needed by the federal government, no other form of taxation, and indeed no other source of revenue of any sort, being of comparable significance. For example during the generation ending in 1913, the combined receipts from customs and excise provided more than 90 per cent of the net ordinary revenue (*i.e.*, excluding loans) of the national government in every year except one.[1]

[1] In 1898 this percentage was reduced on account of a large fortuitous income from sale of public property in the Kansas Pacific Railroad and the Union Pacific Railroad.

The federal tax system was revolutionized by the permanent addition of the income tax — on corporations in 1909 and on individuals in 1913. Shortly thereafter the estate tax, regarded heretofore as a temporary war measure, became a fixed part of the system, when Congress failed to repeal the World War estate tax. The ordinary receipts of the federal government for two recent fiscal years (the twelve months ending June 30) are shown in the following table:[1]

	1929		*1938*	
	Millions of dollars	*Per cent of total*	*Millions of dollars*	*Per cent of total*
Customs	603	14 9	359	5.7
Internal revenue (miscl.)	547	13 5	2,658	42 6
Income and profits tax	2,331	57 7	2,635	42 2
Estate tax	62	1 5	382	6 1
Other ordinary receipts	493	12 2	208	3 3
Total ordinary receipts	4,036	100 0	6 212	100 0

Nothing could show more clearly the revolution in the revenue system of the federal government. Up to 1913, as we have seen, the national government had relied for its revenues almost exclusively upon indirect taxation, leaving the whole field of direct taxes to the states and their local subdivisions. Thereafter the federal government not only invaded the direct tax field; it made direct taxation its chief reliance. In 1929 more than half of the total ordinary receipts came from the income tax. Barely over one-fourth came from the two great indirect taxes, customs and excise, which for a generation before 1913 produced more than 90 per cent of the ordinary receipts.

The depression following 1929 caused a drastic decline in income and excess profits tax receipts, which in 1933 produced only 746 millions, less than one-third of their yield in 1929. This started the pendulum swinging back again toward indirect taxation, though not toward the customs tariff, which continued to decline both absolutely and relatively. It was the excises which came to the Treasury's aid, so much so that in 1938 they were producing nearly five times as much as in 1929 and were even approaching

[1] *Annual Report of the Secretary of the Treasury on the State of the Finances, 1929*, pp. 375 and 420; *1938*, pp 2–3

the relative position they held in the days before the income tax appeared.

The revenue system of the American states. It is out of the question to attempt, in this general study, any detailed account of the tax systems of the forty-eight American states. The following table shows the principal forms of revenue used and the total sums obtained by all the states from each source in the fiscal year 1932.

REVENUE RECEIPTS OF THE STATES, 1932[1]

(In thousands of dollars)

Property tax	323,477
Inheritance tax	143,959
Income tax	47,852
Poll tax	5,321
Licenses and permits	979,385
Other taxes	141,856
Special assessments	26,060
Fines, forfeits, and escheats	5,833
Subventions and grants	249,282
Donations, gifts, and pension assessments	32,034
Highway privileges, rents, and interest	87,996
Earnings of general departments	154,701
Earnings of public service enterprises	10,179
Total	2,207,935

Of the combined state revenue receipts, amounting to 2,208 million dollars, taxation contributed 1,642 millions, or 75 per cent. The heading "licenses and permits" in the table includes the gasoline taxes, motor vehicle licenses, taxes on business, etc. The combined revenue from this group of taxes and licenses furnished 60 per cent of the total tax revenue. The property tax came next, contributing about 20 per cent. The inheritance tax accounted for about 9 per cent, and the income tax, 3 per cent. Of the non-tax revenues the principal item was subventions and grants from other governments, *i.e.*, from the federal government above and from the minor subdivisions of the respective states. The reader will note the small importance of earnings from public service enterprises.

Revenues of American local governments. The following table shows the principal sources of revenue employed by the local sub-

[1] United States Bureau of the Census, *Financial Statistics of States and Local Governments 1932*, 1935, p 7.

divisions of the American states and the amount contributed by each source to the several divisions of local government throughout the country. As we have already learned, the property tax here stands pre-eminent; it furnishes two-thirds of the total revenue receipts of all the local governments and 87 per cent of their total tax revenues.

REVENUE RECEIPTS OF LOCAL GOVERNMENTS, 1932[1]
(In thousands of dollars)

Revenue receipts	*Divisions of government*					
	COUNTIES	CITIES TOWNS VILLAGES AND BOROUGHS	SCHOOL DISTRICTS	TOWNSHIPS	OTHER CIVIL DIVISIONS	TOTAL
Property tax	877,142	2,057,495	1,118,744	195 431	112,195	4,361,307
Licenses and permits	132,050	116,635	184	5,652	2,458	256,979
Other taxes	11,643	77,159	2,108	6,575	126	97,611
Special assessments	32,146	226,430		5,389	31,140	295,105
Subventions and grants	139,077	137,784	317,641	18,830	5,255	618,587
Earnings of general departments	88,956	87,415	63,517	8,535	8,817	257,240
Earnings of public service enterprises	821	487,001	28	2,143	29,826	519,819
Other	31,997	183,675	12,428	5,094	4,140	237,334
Total	1,313,832	3,373,594	1,514,650	247,649	194,257	6,643,982

The United States Census, upon which we are dependent for statistics of state and local government finance, furnishes no information for the states or the combined local governments for years later than 1932. Later statistics are compiled however for the larger cities. The table on the next page for the 94 American cities with population of 100,000 and over for the year 1936 may be of interest.

The reader cannot fail again to note the importance of the property tax. It produced in 1936 over 60 per cent of the total revenues of these 94 large cities; almost exactly 90 per cent of their tax revenues.

[1] United States Bureau of the Census, *Financial Statistics of State and Local Governments 1932*, 1935, p 7.

Revenue Receipts of American Cities with Population of 100,000 and over, 1936[1]

(In thousands of dollars)

Property tax	1,752,378
Poll tax	1,560
Liquor licenses	37,441
Other licenses	162,077
Special assessments and special charges	45,524
Fines, forfeits, and escheats	11,037
Grants in aid	372,683
Donations	2,979
Pension assessments	36,391
Highway privileges	20,094
Rents and interest	84,246
Earnings of general departments	64,493
Earnings of public service enterprises	302,317
Total	2,893,222

Public borrowing: Purposes of borrowing. A government borrows in order to counteract a deficit, realized or anticipated. The circumstances which cause the deficit and so compel the government to borrow are various.

In the absence of extraordinary events which may arise to affect either the expenditure or the income of a government, there is normally an approach toward agreement between the totals on the two sides of the account. Even in the best regulated governments however an exact balance is seldom to be looked for. Moderate deficits and surpluses occur from time to time, and while in the long run they tend to offset each other, the periods of deficit require borrowing to make both ends meet. Such borrowing in order to meet a deficit in the ordinary running expenses is common to all governments. Especially in the United States, where budgetary procedure and financial administration are far from perfect, federal, state, and local governments are frequent borrowers for this purpose. The loans contracted by the federal government to meet the huge deficits of 1931-1938 are discussed in later sections of this chapter.

No nation can meet the extraordinary expenditures of a great war out of its ordinary sources of revenue, however heroically it may expand its tax system and other revenues. Moreover the mag-

[1] United States Bureau of the Census, *Financial Statistics of Cities: 1936*, 1938, pp 58-59.

nitude of war costs is seldom foreseen by the people and their leaders at the war's beginning, realization of the extraordinary contributions which will be required comes slowly, and the war is generally well along, sometimes approaching its close, before the maximum financial effort of which the people are capable is exerted. Often a weak or unintelligent leadership purposely avoids calling upon the people for heavy tax contributions. For these reasons war generally causes large government loans. The American Revolution was almost entirely financed by loans. The cost of the Civil War was about three billion dollars, of which four-fifths was borrowed, the balance being met mainly by taxation. The most gigantic loans in the history of public finance were occasioned by the World War. The United States borrowed about twenty-four billion dollars, or three-fourths of the cost imposed by the war. Great Britain expended forty-two billion dollars on account of the war, obtaining thirty-six billions by means of borrowing Germany, under the guidance of a weak and mistaken leadership, borrowed almost the entire amount of her war expenditure.

Local governments and members of federal governments, like the states of the United States, do not have to finance wars, and since they are seldom confronted with other serious emergencies, emergency borrowing is comparatively rare with them.

When a government engages in an industrial enterprise of any magnitude it usually finds it necessary to borrow the capital required When the Prussian government decided to take over the railway system, it borrowed the money to purchase the existing railroads from their private owners. The United States government borrowed in order to construct the Panama Canal. The state of New York borrowed large sums to enlarge and improve the Erie Canal. American cities occasionally borrow in order to establish water works systems and other municipal industries. Governments also borrow, not infrequently, the money necessary to pay the cost of buildings and other public works needed for the performance of their primary functions. States and towns frequently borrow to build roads and bridges; towns and cities, to build school houses and other public buildings.

The security back of public debts. Government debts differ from ordinary private obligations in that there is usually no pledge of security to guarantee payment. The exception to this statement relates to weak governments with doubtful credit which, when borrowing from foreign capitalists, are sometimes required to pledge certain public property or certain sources of revenue. Such pledges are however the exception, not the rule, and in the case of a strong nation no pledge whatever is ordinarily given beyond the word of the government. In case of default of the ordinary public debt, the creditor has no recourse except either to bring suit in the debtor government's own courts — which is not of much help if the government has deliberately chosen not to pay — or, if a foreigner, to seek to persuade his own government to take diplomatic or, in the last resort, military action against the defaulting government. The history of public finance is replete with cases illustrating all of these principles.

High standing of government obligations. It may appear strange, in the face of the foregoing, that the unsecured debts of sovereign powers generally enjoy such high standing. As a matter of fact, government obligations, especially those of strong national governments, which are almost invariably unsecured, are ordinarily held in high favor by the investing public. Under normal conditions governments can generally borrow at lower rates than are paid by other borrowers, and government obligations are generally quoted on the markets at higher prices in proportion to their yield than any other investments. Bonds of respectable governments stand in the very front frank of "gilt-edged" securities.

The high standing of public obligations rests on certain characteristics of government. In the first place, the life of a government is not limited as is the life of an individual, and it is considered more enduring even than the lives of most corporations. Of chief significance however is the fact that the government's ability to meet its obligations, both interest and principal, is not dependent upon the success of any particular business venture. Most borrowing by individuals and corporations is for investment in business. If the business fails the borrowers may default on the interest or on part or the whole of the principal. The government's ability to

meet its obligations rests upon the taxing power, which may always be exerted to obtain the necessary money provided the debt is not too large. The absence of security is generally counteracted by a most scrupulous regard for its financial obligations, by which most governments seek to protect their credit. In public finance no less than private — indeed even more so — "honesty is the best policy."

It must be confessed that the history of the past generation has tended somewhat to impair the high standing of government obligations. The World War placed upon certain nations a financial strain which was more than they could bear. As a result of budget deficits and reckless issue of paper money, Russia, Germany, and certain other great nations virtually repudiated their earlier public debts. Other nations, not in quite so bad a way as this, nevertheless saw their obligations sink far below par and suffered serious loss of credit; among these were France and Italy. Great Britain and the United States are the two conspicuous examples of great nations which came through that struggle without serious impairment of national credit. Since then monetary depreciation has lowered the value of public debts in many nations. The performance of the United States in 1933 in simultaneously repudiating the gold contract in its national bonds and depreciating the value of the paper money in which alone the bonds may now be redeemed was not favorable to the credit of the United States government. The very size of the debts of some of the leading nations — including the United States — is an element of weakness. Government obligations, though still standing high, do not generally command the confidence which they held a generation ago.

The national debt of the United States. The War of the Revolution was financed principally by loans, and when the new national government of the United States started operations in 1789 it had a heritage of 52 millions of public debt, a sum which appears trivial enough to us of the present generation, accustomed to speak casually in terms of billions, but which the people of that day regarded as a heavy burden upon the shoulders of the new government.

Nevertheless one of the first acts of the government, under the far-sighted leadership of Alexander Hamilton, was the assumption,

as an addition to the national debt, of the debts of the several states incurred in the prosecution of the Revolutionary War. This brought the debt up to 76 millions in 1794. American public opinion, from the very beginning, has looked with disfavor upon a large national debt and has always favored rapid payment. By 1812 the debt of the United States had been reduced to 45 millions. The War of 1812, likewise financed principally by loans, brought the national debt up to 128 millions. Then commenced a period of rapid debt payment, culminating in 1836 in the complete extinguishing of the national debt except for a few thousand dollars, not paid because the holders had not presented the bonds and for whose payment the Treasury had the cash in hand. This represents a unique performance in the records of public finance — a national debt completely paid off.

The country was not long out of debt. Deficiencies in current revenues and the Mexican War restored the debt, which just before the Civil War in 1860 amounted to 65 million dollars. The borrowing of this war provided four-fifths of the three billion dollars which that tremendous struggle cost the United States; on August 31, 1865, the national debt stood at $2,845,000,000. Once more the nation set about the task of debt payment and, by a performance only less remarkable than its previous record, had brought the debt in 1893 down to 839 millions, a reduction of more than two-thirds. Since then there have been numerous debt operations, increases as well as reductions, the principal causes of new borrowing being the resumption of specie payment of the greenbacks in 1879, the currency disorders of 1893–96, the financing of the Spanish War, and the digging of the Panama Canal. In 1916, the year before the United States entered the World War, the national debt amounted to about a billion and a quarter. The World War was financed to the extent of three-fourths of its cost to the United States by borrowing, principally by the issue of the Liberty loans, Treasury notes, certificates of indebtedness, and war savings certificates. The total borrowing on account of this war makes everything preceding pale into relative insignificance; on August 31, 1919, the national debt amounted to $26,596,701,648. Once again the traditional American policy of debt reduction was vigorously

pursued. Large appropriations for payment of the debt were included in the regular annual budgets; in addition substantial surplus revenues were regularly available for debt retirement, and the debt by December 31, 1930, had been reduced to 16,026 millions, a reduction of two-fifths of the debt in eleven years. The subsequent growth of the national debt will be noted in connection with our examination of "New Deal" finance later in this chapter.

Government surplus. The mere mention of such a thing as a government surplus in this day and generation may possibly be taken as indicating a deficient sense of humor And yet there have been government surpluses, and they have presented some interesting problems Indeed it is not too much to say that, as compared with the moderate routine deficit in operating revenue, a sizable surplus is apt to be the more difficult to handle. In private finance a surplus is of all things the most fervently to be desired, faced with a glittering array of possibilities for its disposal, the idea that it may become an embarrassment is the last to enter the mind of the fortunate one But we are here confronted with one of the several important differences between public and private finance Suppose that a sovereign government (the only kind of government which, with proper financial management, should ever have a formidable surplus) finds itself, on account of some unexpected accretion of revenue or simply as the result of bad budgeting, in the possession of a large surplus What, let us ask, shall be done with it? Simply to store money in the vaults of the treasury, withdrawing a material part of the country's money from circulation, is not permissible. Large idle deposits in the banks are little better. Investment naturally occurs; this is the obvious solution of the problem in private finance But the spectacle of a democratic government announcing itself as a would-be purchaser of the stocks and bonds of business corporations or going quietly into the market and purchasing securities presents a prospect of favoritism, corruption, and public loss which will be easily visualized May not the surplus be used to reduce the public debt? Here is indeed the one safe outlet, and so long as surplus revenues are devoted to retirement of debt no serious consequences will ensue, and the result will be wholly favorable. Unfortunately the presence of surplus money offers so

great a temptation to the politicians in control of the government's finances that it is not certain that the safety-valve of debt payment will always be permitted to operate to the full extent.

The United States has experienced some interesting encounters with surplus revenues, which illustrate these principles. Beginning immediately after the close of the War of 1812, the growing customs revenue together with fortunate sales of public lands brought receipts in excess of the government's needs, so that for twenty years there was an almost uninterrupted condition of surplus revenue. Until 1836 the public debt furnished the ready means of disposal, and no serious embarrassment was felt. But in that year the debt had been extinguished; and the surplus continued. Congress and the people were in a quandary, and after a long wrangle, the entertaining details of which we must pass by, it was decided to divide the surplus among the states. An outright gift appearing likely to be held unconstitutional, the money was transferred to the states as a loan, with the tacit understanding that its repayment would never be demanded. The problem was thus passed on to the states, which disposed of the money in a variety of ways but generally without any great permanent profit. Among other consequences, the surplus and its disposition were partly responsible for the disastrous panic of 1837, which incidentally reversed the government's problem by quickly turning the surplus into an embarrassing deficit.

The next period of surplus culminated in the decade 1880–1890. In August, 1888, the government's surplus was equal to one-fifth of the country's entire monetary circulation. Now again the public debt offered a convenient outlet, and we have already noted the extraordinary record of debt payment which came to an end in this decade. However it did not prove feasible to use the entire surplus in this way, and the tempting hoard made political fingers itch. Congress, incited by the administration, proceeded to dissipate the surplus in a riot of extravagant expenditures, for pensions, river and harbor improvements, and public buildings, setting a standard from which the taxpayers suffered for years thereafter. There have been occasions of surplus revenue since, notably in 1907 just before the panic of that year, and these surpluses have always been trouble-

some. Historically it appears that in public finance a surplus presents so embarrassing a problem that few democratic governments are able to cope with it.

The budget. Under sovereign constitutional governments all taxes and other revenues and all expenditures must be authorized by law. This requires that at certain stated intervals the law-making body, be it Parliament, Congress, state legislature, or other law-making authority, must take account of the state of the public finances, determine what services shall be performed by the government, estimate the probable cost of such services, and provide taxes and other income to meet the anticipated expenditures. In determining its fiscal policy, the legislature will find it helpful to take account of the income and expenditure of the past year or recent years. Such account, together with the estimates and plans for the coming year, is called the *budget*.

The subordinate grades of government — counties, cities, towns, school districts, etc. — also have use for the budget. Although they have no sovereign power, they have been delegated a considerable authority over their own finances, under which the local legislative body — board of aldermen, city council, town selectmen, or whatever it may be — determines the expenditures and at least the rates of taxation for the district.

Budget procedure has been brought to a high degree of perfection by certain governments, the most notable being the British Parliament. In the United States the budget has been generally more honored in the breach than in the observance by all grades of government. Its efficient development has been hindered in the federal government by the constitutional separation between the executive and legislative arms of government and the bicameral Congress, features which have been generally carried down into the political organization of the states; and there have been other contributing causes of the poor record in America. Only since the World War has the federal government developed any semblance of a budgetary procedure, and, barring a few notable exceptions, the states and municipalities have been no less backward. Even now the United States budget falls far short of the standards set either by theory or by the practice of such nations as Great Britain.

Financial aspects of the "New Deal": General character. The crisis of 1929 and the succeeding period of depression witnessed a governmental program in the United States which embraces many projects of interest to the student of economics. The more important of these projects are examined at various places in this book in accordance with their subject matter. The present appears to be the proper place to investigate the important financial aspects of the "New Deal," as this program came to be called.

In its attitude toward finance, the philosophy of the "New Deal" presented two or three faces, distinct, and not free from contradiction. First and most obvious, the operation of the several parts of the program occasioned the expenditure of enormous sums of money. A reckoning of this account appears later in this chapter. On the other hand, there was apparent an inclination, evidently shared by the administration, to regard spending as something more than this. Government spending was considered desirable in itself, as a means of "putting money in circulation," stimulating purchasing power, "priming the pump," and restoring normal industrial activity. "Spending ourselves out of depression" was one face of the "New Deal" philosophy. From this viewpoint, all government spending carried a favorable presumption, tending to brush aside traditional notions of economy and to relegate to the background consideration of the balanced budget.

The government was, at the start at least, not wholly unmindful of economy. One of the first acts of the Congress that assembled in special session in March, 1933, was the enactment of the economy act, restricting drastically the payment of veterans' pensions and benefits, and reducing the salaries of the employees of the federal government. Even such modest service to economy was however not permitted by Congress to continue more than a year. In March, 1934, over the veto of President Roosevelt, Congress restored three-fourths of the reduction in veterans' compensation made by the act of 1933 and also dictated gradual restoration within the coming year of the 15 per cent pay cut which had been made in federal salaries. In 1936 Congress granted another extraordinary

veterans' bonus, which added 1½ billions to the expenditures of the fiscal year 1936 and another half billion in 1937.

The double budget. Inconsistency between the claims of economy and the lavish expenditures of the "New Deal" was officially disavowed through the device of the double budget, by which means the administration sought to place special emphasis upon the "emergency" character of the "New Deal" expenditures, in distinction from those expenditures which might be presumed sufficient to meet the ordinary needs of the government. The enormous deficit in government operations was after the fiscal year 1933 relegated to the "emergency budget," while the administration continued to proclaim its interest in economy and the preservation of the public credit.

The double budget is not an invention of the "New Deal." It will be recalled as a device employed by France during the long reconstruction period following the World War. Of course the double budget has no magic power to banish a deficit. The deficit, as disclosed by the total account of revenues and cost payments or by the mounting public debt, goes on unhindered. In its basic implication moreover the "New Deal" budget was from the start misleading, through inclusion in the emergency budget of considerable expenditure which was destined — and even intended — to be permanent, thus making the balance between revenues and general expenditures essentially fictitious.

Public revenues and expenditures. On the revenue side of the picture the "New Deal" has offered nothing of great moment, and no substantial effort has been made to provide taxation to meet the extraordinary cost of the existing activities of government. The revenue act of May, 1934, was announced as aimed chiefly at "tax avoidance." It was estimated to be capable of adding 417 million dollars to the annual revenues when fully in operation, obviously only a small fraction of current deficits. The revenue act of 1935 was obviously intended chiefly as a means toward redistribution of wealth; the President, in January, 1936, estimated its yield at only 222 millions. The revenue act of 1936 undertook to provide new tax revenue to make good losses occasioned by the invalidating of the processing taxes on agricultural products by the Supreme Court

and to meet the cost of the soldiers' bonus legislation, together with changes in the corporation taxes and the individual income tax calculated to raise 620 million dollars a year. Similarly the federal revenue acts of 1937 and 1938 sought chiefly amendment in details and improvement of administration and disclaimed any intention to increase materially the demands of taxation. It has clearly been the policy of the administration generally to avoid increased taxation and to rely upon credit for meeting the "emergency" expenditures.

It is not necessary at this point to inquire into the details of the public expenditure of the "New Deal." Money was spent, in huge sums, for various purposes; the most important of these are discussed at appropriate places in other chapters of this book. The total expenditure for "recovery and relief," as officially reported, during the fiscal years 1931–1938 (*i.e.*, up to June 30, 1938) was 18,390 million dollars. If to this we add 2,741 millions, the amount officially estimated for 1939,[1] and 2,019 millions as officially estimated for 1940,[1] we reach the total of 23,150 millions to June 30, 1940. What chiefly concerns us here is the total cost of government and its relation to revenues during the years of depression and "New Deal" financing. For that we may proceed to the discussion in the following section.

Depression, "New Deal," and the federal finances. In order to visualize clearly the net effect of depression and "New Deal" upon the income account of the United States government, we may profitably make our start at the beginning of the fiscal year 1929, thus including for the sake of comparison the last two years in which prosperity still smiled upon the National Treasury. The table on the next page shows the revenue receipts and the cost payments (*i.e.*, not including the proceeds of borrowing or sums expended in redemption of debt) for twelve years, the last two years as estimated in the *Budget of the United States Government for the Fiscal Year 1940*.[2]

[1] *Budget of the United States Government for the Fiscal Year 1940.* It is to be noted that owing to changes in the official classification these figures for 1939 and 1940 are not on the same basis as the previous figures for "recovery and relief."

[2] For facts such as these relating to the finances of the United States government, the reader is referred to *Annual Report of the Secretary of the Treasury on the State of the Finances* (known as the "Finance Report"), *Budget of the United States Government* (annual), and *Daily Statement of the United States Treasury.*

REVENUE RECEIPTS AND COST PAYMENTS, UNITED STATES GOVERNMENT
(In millions of dollars)

Year	Revenue Receipts	Cost Payments			Surplus (+) or Deficit (−)
		GENERAL	RECOVERY AND RELIEF	TOTAL	
1929	4,033			3,299	+734
1930	4,178			3,440	+738
1931	3,190			3,652	−462
1932	2,006	3,848	893	4,741	−2,735
1933	2,080	3,404	1,277	4,681	−2,602
1934	3,116	2,743	4,002	6,745	−3,630
1935	3,800	3,145	3,657	6,802	−3,002
1936	4,116	5,186	3,291	8,477	−4,361
1937	5,294	5,165	2,836	8,001	−2,707
1938	6,242	5,267	2,434	7,701	−1,459
1939	5,520	6,751	2,741	9,492	−3,972
1940	5,669	6,976	2,019	8,996	−3,326

Attention should be called to the fact that the expenditures here shown are less than they usually appear in the official statements (general and special accounts), owing to the exclusion of sums (averaging a little less than half a billion dollars a year through 1936 and about 90 millions in the last four years) expended for reduction of debt; such items are not true cost payments. It should also be noted that the receipts and payments for 1934, as here shown, do not take account of $2,811,375,757 "increment resulting from reduction in the weight of the gold dollar," of which 2 billions were devoted to the "exchange stabilization fund" Transactions in the gold increment and the stabilization fund (comparatively small in amount) are excluded likewise from the figures for later years [1]

On the revenue side of this account, the blighting effect of depression is clearly discernible in the decline of receipts in 1932 and 1933 to only about half the amounts received in 1929 and 1930 The

[1] Trust fund transactions are, properly excluded from the figures for the years 1931 39 Owing to the form of the official reports, it is not possible to exclude these transactions from the figures for 1929 and 1930 The differences are not sufficient to affect the comparisons here made Refinements such as these account for discrepancies between certain figures here and the corresponding ones in the table of expenditures of the United States government appearing earlier in this chapter.

slight rise in 1933 and the more substantial gains through 1938 are in consequence of tax increases plus some improvement in general economic conditions. The very substantial reduction in revenue predicted for 1939 and 1940 reflects the depressed condition of national industry which prevailed in the fall of 1937 and through the calendar year 1938.

Efforts to fight depression by public spending were undertaken by the Hoover Administration, with consequences as shown by the figures of 1931, 1932, and 1933. Late in the fiscal year 1933 the expenditures of the "New Deal" proper began to have their effect. It will be noted that total expenditures rose to about $8\frac{1}{2}$ billions in 1936 and then declined slightly in the next two years. The estimates of nearly $9\frac{1}{2}$ billions for 1939 and about 9 billions for 1940 are consequent upon the "pump-priming" policy employed to meet the return of depression conditions in the fall of 1937.

The handsome surpluses of 1929 and 1930 are typical of the period following the World War, during which the National Treasury enjoyed great prosperity. In the fiscal year 1931 there began a series of heavy deficits, due for the first three years in part to anti-depression expenditures but principally to declining revenues, and mainly chargeable to "New Deal" costs in 1934 and thereafter. The fiscal year 1938 came nearer to a balance than any previous year since 1931, but the official estimate for 1939 (made in January, 1939) shows a deficit greater than in any year except 1936, and the estimated deficit for 1940 is only slightly less.

"General and emergency" costs. In the foregoing table the total governmental costs have been separated into "general" costs and expenditures for "recovery and relief," or "emergency" costs. This separation is based upon the official reports of the Treasury and the Budget Bureau. The apparent rapid decline of expenditures for recovery and relief since 1934 is on the whole not real, being chiefly the result of transferring to the general budget the cost of various agencies, such as the Tennessee Valley Authority, the Agricultural Adjustment Administration, the Civilian Conservation Corps, and the Farm Credit Administration, which were originally in the emergency budget. This accounting shift shows the misleading nature of the apparent balance of the "general" budget

in the years 1934 and 1935. It shows also the lack of foundation for the inference that, upon the termination of the "emergency," the total cost of the United States government may be expected to decline to anything like the previous level.

In the 1940 budget a new classification of expenditures was set up. Expenditures are now divided into two main groups, "ordinary" and "extraordinary," the latter including public works, loans, subscriptions to stock, etc. to various government agencies, and the cost of a "new national defense program" (for 1940) in addition to the now more limited items of "unemployment relief." The following figures for 1939 may be compared with the next to last line in the preceding table:

Ordinary expenditures	$5,251
Public works	1,229
Unemployment relief	2,741
Loans, subscriptions to stock, etc (net)	271
Total	9,492

Progressive increase in cost of government. After every period of extraordinarily swollen public expenditures, the United States has found itself upon a new permanent level of governmental costs far above that which formerly prevailed. To the time of the Civil War, the annual per capita cost of the United States government was never much above two dollars; following the Civil War it never fell below four dollars. In the fiscal year 1916, the total ordinary expenditures of the national government were 734 million dollars; since the World War they have never been below three billions, except in 1927 (when the amount was 2,974 millions). The charge for interest alone in 1920 (1,020 million dollars) exceeded by one-third the entire cost of the government in 1916 or any previous year.

That the aftermath of the existing régime of swollen expenditures and deficit financing will run counter to past experience is hardly to be expected. With a debt greater than ever before, the government will be bearing a charge for interest alone well over a billion dollars. Assuming that the American people adhere to their traditional policy of debt payment after every great emergency, there will be an appropriation for debt payment each year which may be

as much as half a billion (on the precedent of the years following the World War). That the extraordinarily expanded governmental machine which now exists can ever be reduced to its pre-depression dimensions appears most unlikely in the light of past experience. The year 1939 showed the highest level of expenditure ever experienced by the federal government in time of peace. 1940 will present at least a close second to this record.

Borrowing to meet the cost. The gross debt of the United States stood at its then maximum of 26,597 million dollars on August 31, 1919. It reached its low point since the World War on December 31, 1930, when the debt amounted to 16,026 million dollars. The deficits of the depression years reversed the progress of debt payment, and by June 30, 1934, all of the gain since the World War had been lost. Succeeding deficits brought the debt to the unprecedented total of 40,773 million dollars on August 10, 1939. Thus have been broken all records for the United States, with the national debt exceeding by more than 14 billions its maximum height as occasioned by the World War. It appears certain that the debt will continue to increase for at least another year. Official estimates indicate a total of over forty-four billions on June 30, 1940.

SUGGESTIONS FOR FURTHER READING ON PART V, TAXATION AND THE BUSINESS OF GOVERNMENT

I PUBLIC FINANCE. GENERAL TEXTS

LUTZ, H. L. *Public Finance* Third edition New York, 1936
JENSEN, J. P. *Government Finance*. New York, 1937
KING, C. L *Public Finance*. New York, 1935
SHULTZ, W. J. *American Public Finance* Second revised edition. New York, 1937
BUEHLER, A G. *Public Finance* New York, 1936
BASTABLE, C. F. *Public Finance* Third edition. London, 1903
SHIRRAS, G. F. *The Science of Public Finance*. Third edition London, 1936
PLEHN, C. C. *Introduction to Public Finance*. Fifth edition. New York, 1926
ADAMS, H. C. *The Science of Finance*. New York, 1898
SELIGMAN, E R. A. *Studies in Public Finance*. New York, 1925
PIGOU, A. C. *A Study in Public Finance*. New York, 1928
DALTON, H *Principles of Public Finance*. Ninth edition. London, 1936
BULLOCK, C. J. (Editor). *Selected Readings in Public Finance*. Third edition. Boston, 1924

Mills, M. C., and Starr, C. W. *Readings in Public Finance and Taxation.* New York, 1932

Fagan, E. D., and Macy, C. W. *Public Finance, Selected Readings.* New York, 1934

Dewey, D. R. *Financial History of the United States.* Twelfth edition. New York, 1934

Shultz, W. J., and Caine, M. R. *Financial Development of the United States.* New York, 1937

Noyes, A. D. *Forty Years of American Finance.* New York, 1909

—— *The War Period of American Finance.* New York, 1926

II. Taxation

Green, W. R. *The Theory and Practice of Modern Taxation.* Chicago, 1933

Wells, D. A. *Theory and Practice of Taxation.* New York, 1911

Armitage-Smith, G. *Principles and Methods of Taxation* New York, 1906

Stamp, J. *The Fundamental Principles of Taxation.* New York, 1929

Lyon, W. H. *Principles of Taxation.* Boston, 1914

Brown, H G. *The Economics of Taxation.* New York, 1924

National Industrial Conference Board, Inc. *Taxation and National Income.* New York, 1922

—— *Tax Burdens and Exemptions* New York, 1923

—— *Tax Burdens and Public Expenditures.* New York, 1925

—— *Cost of Government in the United States, 1935 1937.* New York, 1938

—— *Cost of Government, 1923 1934.* New York, 1934

United States Bureau of the Census *Financial Statistics of State and Local Governments · 1932.* Washington, 1935

Tax Research Foundation. *Federal and State Tax Systems.* Chicago, Annual

Twentieth Century Fund *Facing the Tax Problem.* New York, 1937

Shoup, C., and others (Twentieth Century Fund). *Studies in Current Tax Problems.* New York, 1937

Brown, H. G. *The Economic Basis of Tax Reform* Columbia, 1932

Seligman, E R. A. *Essays in Taxation* Ninth edition. New York, 1921

Comstock, A. *Taxation in the Modern State* New York, 1929

Ely, R T. *Taxation in American States and Cities.* New York, 1888

Howe, F. C. *Taxation and Taxes in the United States under the Internal Revenue System, 1791 1895* New York, 1896

Seligman, E. R A. *The Income Tax.* New York, 1911

Spaulding, H B. *The Income Tax in Great Britain and the United States.* London, 1927

Buehler, A. G. *The Undistributed Profits Tax.* New York, 1937

Kendrick, M. S. *The Undistributed Profits Tax* Washington, Brookings Institution, 1937

Leland, S. E. *The Classified Property Tax in the United States* Boston, 1928

Jensen, J P *Property Taxation in the United States* Chicago, 1931

National Industrial Conference Board, Inc. *State and Local Taxation of Property.* New York, 1930

West, M *The Inheritance Tax.* New York, 1908

Shultz, W. J. *The Taxation of Inheritance.* Boston, 1926
Report of the National Committee on Inheritance Taxation. Washington, 1926
Hardy, A. *Inheritance and Other Like Taxes.* New York, 1929
Buehler, A. D. *General Sales Taxation.* New York, 1932
National Industrial Conference Board, Inc. *General Sales or Turnover Taxation.* New York, 1929
Haig, R. M., and Shoup, C. S. *The Sales Tax in the American States.* New York, 1934
National Industrial Conference Board, Inc. *State and Local Taxation of Business Corporations.* New York, 1930
Alvord, E. C., and others. (Tax Policy League, Symposium.) *How Shall Business be Taxed?* New York, 1937
Fairchild, F. R., and associates *Forest Taxation in the United States* (Report of the United States Forest Taxation Inquiry) Washington, 1935
Studenski, P., and others. *Taxation and Public Policy.* New York, 1936
Peck, H. W. *Taxation and Welfare.* New York, 1925
Fillebrown, C. B *The A B C of Taxation.* New York, 1909
George, H. *Progress and Poverty.* Numerous editions.
Post, L. F. *The Taxation of Land Values.* Fifth edition. Indianapolis, 1915
Hobson, J. A *Taxation in the New State.* New York, 1920
Lutz, H. L. *The State Tax Commission.* Cambridge, 1918
National Tax Association *Proceedings of the National Conference* — Annual from 1907 *Bulletin*, Monthly from February, 1916

III. Budget, Government Borrowing, War Finance, Etc.

Gulst, H W *Public Expenditures.* New York, 1928
Love, R A *Federal Financing* New York, 1931
Stourm, R. *The Budget* (English translation) New York, 1917
Cleveland, F A *The Budget and Responsible Government.* New York, 1920
Buck, A. E *Budget-Making.* New York, 1921
Plehn, C. C. *Government Finance in the United States* Chicago, 1915
Willoughby, W F. *The National Budget System.* Baltimore, 1927
Cleveland, F. A. *Municipal Administration and Accounting.* New York, 1919
Upson, L. D. *Practice of Municipal Administration.* New York, 1926
Buck, A. E., and others. *Municipal Finance.* New York, 1926
Grice, J. W. *National and Local Finance.* London, 1910
Adams, H. C. *Public Debts.* New York, 1887
Stewart, P. W., and Tucker, R. S. (Twentieth Century Fund) *The National Debt and Government Credit: Factual Findings.* New York, 1937
Hollander, J. H. *War Borrowing.* New York, 1919
Noyes, A. D. *Financial Chapters of the War.* New York, 1916
Grady, H. F *British War Finance, 1914–1919.* New York, 1927
Jèze, G. P. A., and Truchy, H. *The War Finance of France.* New Haven, 1927
Bergmann, C. *The History of Reparations.* Boston, 1927

MOULTON, H. G., and PASVOLSKY, L. *War Debts and World Prosperity.* Washington, 1932

DEXTER, P., and SEDGWICK, I. H. *The War Debts.* New York, 1928

SERING, M. *Germany under the Dawes Plan.* London, 1929

BOGART, E. L. *Direct and Indirect Costs of the Great World War.* New York, 1919

BASS, J. F., and MOULTON, H. G. *America and the Balance Sheet of Europe.* New York, 1921

World Peace Foundation. *Pamphlets on Reparations.* Boston

PART VI

INTERREGIONAL TRADE AND EXCHANGE

XXXIII

INTERREGIONAL TRADE

The bases of interregional trade: Fundamental motive. Trade, in its simplest form, consists in the exchange of goods in the same place. Most trade however involves more; namely, the exchange of goods between different places. It is this sort of trade which presents the most interesting and important economic problems, and since the principles of the simpler form of trade are included in those of the more complex, we are justified in confining our further study of trade to the subject of interregional and international trade

So many popular fallacies have clustered about the problem of interregional trade that it is essential to inquire somewhat minutely into the nature of the gain that comes from regional division of labor and interregional trade. The trade in some articles is readily enough explained by the fact that in no other way could the community possibly obtain these particular goods. Diamonds, emeralds, and certain other precious stones could not be enjoyed by Americans if they were not brought in through trade. The residents of a great industrial city could not have fresh fruits and vegetables save for the trade of the agricultural hinterland.

The more complex problems of trade however have to do with the many commodities which apparently might be produced at home and which are nevertheless regularly brought in from other regions. The people of New York State could make their own automobiles instead of leaning so heavily upon Michigan; the United States could produce all the sugar it needs instead of only a part. In fact the literal impossibility of producing an article at home can seldom be cited as the reason for its import. The real reason is that, by calling upon the resources and industry of other communities, goods may be obtained at lower cost than would be incurred in their production at home. Each region thus looks for the things it needs to those parts of the country or of the world where each com-

modity is produced most cheaply and conversely finds the rest of the country and of the world calling upon it for those particular commodities which it produces at lowest cost. So each region, because of trade, is enabled to devote its resources and energies to the lines of production for which it is best adapted as evidenced by its ability to produce at low cost.

Local advantages and their causes. Many factors combine to determine whether one region can produce a good more cheaply than another. The rich soil and favorable climate of Cuba and Hawaii make it possible to grow sugar at lower cost than in the continental United States; the exceptional quality of the manganese ore of the Caucasus Mountains, India, and Brazil gives to those regions such an advantage that virtually all the manganese used in the steel industry of the United States is imported in spite of the presence here of large reserves of low-grade manganiferous iron ore. In other cases the low cost of production is due to the economies which arise from large-scale production, the extensive use of machinery, and the minute subdivision of operations. The great reserves of iron and coal of England, together with an intelligent and industrious population, large accumulations of capital for investment, and the application of large-scale methods of production, have given the British advantages in the manufacture of many articles. Numerous other illustrations of the way in which the products of a region have been determined by superiority in natural resources, by peculiar skill on the part of its laborers, by relatively great accumulations of capital, or by economies in the organization of industry will at once occur to the reader. The essential fact is that differences in productive advantages are enjoyed by different regions and that these lead to regional specialization in production and the exchange of the surplus products through trade.

The gain from specialization and trade. It is doubtless evident from what has been established to this point that specialization by each region in the production of the goods for which it is best fitted must increase enormously the total of the world's production and correspondingly add to the goods available for the satisfaction of human wants, but the situation will be clearer and the causal relation brought more out into the open if we inquire somewhat

minutely into the economic effects of interregional trade. For this purpose we may set up a hypothetical case which, in spite of its obvious artificiality, will serve to expose the principle in which we are interested.

Let us suppose that country A and country B can produce either wheat or corn, but that A has an advantage in the production of wheat, while B has an advantage in the production of corn. To make the example concrete, let us say that with the expenditure of a unit of labor and capital (composed of so many hours of labor, so much machinery, fertilizer, etc.) A can produce either two bushels of wheat or one bushel of corn, while B with the same expenditure of effort (one unit of labor and capital) can produce either one bushel of wheat or two bushels of corn. Then if A and B each spends two units of labor and capital and each produces both corn and wheat the result will be:

A	2 bu. of wheat + 1 bu. of corn
B	1 bu. of wheat + 2 bu. of corn
Total	3 bu. of wheat + 3 bu. of corn

If each country spends its two units in the production of the good in which it has a real advantage the result will be as follows:

A	4 bu. of wheat
B	4 bu. of corn
Total	4 bu. of wheat + 4 bu. of corn

In this case there is a gain in total product of one bushel of wheat and one bushel of corn, but neither country wants to be without the other grain. Rather than do without corn entirely, A would be willing to spend one unit in producing a bushel. Let us then further assume that A exchanges wheat for corn with B at the rate of one bushel of wheat for one bushel of corn.[1] If two bushels are sold, each country would have wheat and corn, as follows:

A	2 bushels of wheat + 2 bushels of corn
B	2 bushels of wheat + 2 bushels of corn

Comparing this table with the first one, we see that A has gained a bushel of corn and B has gained a bushel of wheat as the result of specializing in production and exchanging the surplus.

[1] In these illustrations we ignore the transportation charges.

It may however happen that a country finds itself in the unhappy position of having no natural advantages. Its soil is sterile, nature has not endowed it with any mineral resources of prime importance, and its population is lethargic; yet we find the country has an export trade. On the other hand, there are countries which are so favorably situated that they have advantages in many different lines of production; yet they regularly import certain goods from other countries which are actually inferior to them in those lines. This situation, as to the practical truth of which there can be no possible doubt, may seem anomalous if our previous example has led us to the erroneous conclusion that each country will necessarily produce for itself everything in the production of which it has an advantage, with the corollary that the country which has no superiority in any line can have no production at all.

To explain the situation let us assume two other countries, C and D. Country C with one unit of labor and capital can produce either fifteen yards of cloth or one ton of steel. Country D with one unit of labor and capital can produce ten yards of cloth or half a ton of steel. C has an advantage in the production of both cloth and steel, but its advantage in the production of steel is greater. Compared with C, D has no advantage in the production of either commodity, but its "least comparative disadvantage" lies in the production of cloth. If each country devotes one unit of labor and capital to the production of cloth and two units of labor and capital to the production of steel, the amounts produced by each will be as follows:

C with 3 units	15 yards of cloth + 2 tons of steel
D with 3 units	10 yards of cloth + 1 ton of steel
Total	25 yards of cloth + 3 tons of steel

If however with the same expenditure of effort as before C specializes in the production of steel, in which its comparative advantage is greatest, and D specializes in the production of cloth, in which its comparative disadvantage is least, the amounts will be as follows:

C with 3 units	3 tons of steel
D with 3 units	30 yards of cloth
Total	30 yards of cloth + 3 tons of steel

From the standpoint of total production this specialization has brought about a gain of five yards of cloth, but this is of no particular interest to either country unless the terms of exchange are such that each country can gain thereby. C would be unwilling to sell steel for less than fifteen yards of cloth a ton, since it would be more profitable to manufacture the cloth at home; D, for the same reason, would be unwilling to pay more than twenty yards of cloth for a ton of steel. A rate of eighteen yards of cloth for one ton of steel falls within these limits, and we will assume that this is fixed and that C sells D a ton of steel for eighteen yards of cloth. Then we find, after the exchange has been affected, the following situation.

C	18 yards of cloth + 2 tons of steel
D	12 yards of cloth + 1 ton of steel

Comparison of this with the first table will show that C has gained three yards of cloth while D has gained two yards of cloth. In spite of C's greater productive advantage in both industries, it is worth its while to specialize and exchange products with D. Of course the people of D still have less to show for their efforts than the people of C, but they would have even less if they were unable to specialize on cloth and get their steel by exchange.

The law of comparative advantage. From this example we derive an important generalization: it is to the advantage of a country or region to devote its energies, not to all the lines of production in which it may have superiority, but to those in which its superiority is greatest, provided trade gives it the opportunity to secure other goods from abroad. Likewise the country which has no line of superiority will find it advantageous to devote its energies to those lines in which its inferiority is least marked, provided the opportunity of trade with other regions is open to it. This generalization is called the *law of comparative advantages* and may be stated as follows: *A country gains by concentrating its energies in the production of those goods in which it has the greatest comparative advantages or in which it has the least comparative disadvantages.*

Our conclusion is that regional specialization and interregional trade promote economy in the use of the productive forces of society. By utilizing the most fertile soils and the richest mines, by

turning labor and capital into the industries for which each region is best suited, society as a whole has a greater quantity of consumable goods to share among its members, and each region is enabled to secure a greater quantity of consumable goods than otherwise would be possible.

Since the gain from trade consists in the increased command over consumable goods, it follows that imports are the aim and end of commerce. We export in order that we may import. This conclusion is somewhat at variance with popular belief, but if it be true that commerce exists to serve men's wants, its benefits are to be measured by the command it gives over consumable goods. Our exports are of importance only in so far as they provide an easier means of securing the goods imported than by making them at home, and the greater the quantity of imports which we can obtain with a given quantity of exports, the more fortunate our position and the greater the national income. A falling off in the export trade is to be considered just cause for public alarm only because it may indicate that our buying power in foreign markets has decreased and therefore our command over foreign goods is less than formerly.

These conclusions were derived without reference to political boundaries, and they apply as well to regional specialization within the confines of a single country as to specialization between countries which are politically distinct. The gain which comes to the United States from purchasing the fine woollens of England with wheat is of exactly the same character as that which the people of New England experience when they buy western foodstuffs with their manufactured goods.

Diminishing returns and the law of comparative advantage. In every country hundreds and thousands of different articles are produced. Why, if the law of comparative advantage really holds good, should not each country, or each section of each country, concentrate on the production of one article only, and secure all the other goods it needs by exchange?

In partial answer to this question it may be said that the existence of a comparative advantage in the production of any commodity depends to some extent upon the quantity produced. There is a

limit to the quantity of any good which a community can produce without an increase in unit cost; sooner or later all industry operates under conditions of increasing cost. A rising marginal cost of production is an indication that the comparative advantage in the production of additional units of the good is becoming less. The United States for example has such decided advantages in the production of wheat that it regularly has a surplus of several hundred millions of bushels for export, but if it were to attempt to supply the entire world with wheat the cost per bushel would rise far above that which would obtain in other wheat-producing countries or even in some at least of the countries which now import wheat. Rather than pay the price which in such circumstances would be necessary to remunerate the American farmers for their efforts, foreigners would buy wheat from other sources or grow wheat at home. The law of diminishing returns therefore sets limits to specialization, and we find few products of any consequence in the world's commerce which are produced in one region exclusively, just as we find few important regions with only one industry.

Transportation and the law of comparative advantage. To complete the answer to the question raised above we must also take into account the factor of transportation In our hypothetical examples we ignored the costs of transportation, but in actual life they offer one of the greatest obstacles to the development of regional specialization and trade The reason that the merchants of any region import a commodity from outside is because it can be sold in the domestic market at a price which is lower than the price at which a similar article of domestic manufacture can be sold, and in determining the sale price the cost of transportation must be included. English manufacturers may be able to produce certain grades of brick at a lower cost than can American manufacturers, but if the cost of shipping them to New York is so great as to raise their price above that of similar domestic bricks, the advantage of the English manufacturers so far as the New York market is concerned is without significance. There are many goods of which this is true, goods which are commonly produced in almost every locality, because the cost of transporting them from one place to another is so great as to wipe out entirely the advantage which any

one locality may have. Nearness to market therefore gives an advantage in production which may outweigh positive disadvantages, as is shown for example by the fact that common bricks are almost everywhere a local product, seldom brought in from any great distance, even though other places may enjoy advantages in their manufacture.

The control of trade. Every producer of goods naturally seeks to sell his product in the market where he obtains the highest price, be it near at hand or far away. The consumers just as naturally endeavor to make their purchases where goods may be obtained most cheaply. All the agencies of trade and transportation are enlisted in the effort to aid producers and consumers in the accomplishment of their respective purposes; every class of middlemen is actuated by the same motives as the other producers and the consumers and strives to buy in the cheap market and sell in the dear market. There is a tendency therefore for goods to move from where their prices are low to where their prices are high. By thus reducing the supply in one market and increasing it in another, the price in the one is raised and in the other is lowered, till the prices in different markets are brought into equilibrium.

Equilibrium of prices does not mean equality of prices. The movement of any commodity from the low priced to the high priced market reaches a limit when, as the result of such movement, the difference between the two prices is brought down to the cost of shipping the commodity from the first market to the second. Thereupon the movement ceases, price equilibrium having been reached. If transportation could take place instantaneously and without cost we would find the price of each commodity the same in all markets. As it is, equilibrium is attained when the prices of a commodity do not differ by more than the cost of transportation. This principle may be stated in the following law: *In the absence of restraints on trade, the prices of any standardized commodity in two markets tend to differ by an amount not greater than the cost of transporting the commodity from the market of lower price to that of higher price.*

Thus price and the cost of transportation control the flow of goods from one place to another. There will be trade so long as

there would otherwise be a difference in the prices of the same commodity in two markets greater than the cost of transportation between them. Where the difference is less than this there is no interregional trade, as is the case of many commodities of which the supplies are relatively local; bricks for example. The ability to specialize to advantage is evidenced by a relatively low cost of production, which enables the producers of one region to undersell those of other regions, and the opportunities for the extension of regional specialization and trade are brought to light by differences in the prices of the commodity in the various markets. Under competitive conditions price serves as an automatic control over regional division of labor and interregional trade.

The restraints which may prevent trade and so prevent the attainment of price equilibrium are numerous. Frequent examples have already been encountered in our study, and we shall later give attention to one of the most important, the customs tariff.

The case of labor. In the present connection we must take special note of one commodity (if this term may be so used) which is peculiarly susceptible to obstacles restraining movement in response to inequalities of price. This is labor, whose price — wages — may vary widely and permanently between different countries and different parts of the same country without being corrected by the migration of the laboring populations. Domestic servants in one part of the world may be obtained at an extremely low figure, while their services in another country command a high price. Labor of a particular type is paid different rates in different sections of the country, and there is often material difference between the wage scales of the city and the immediately surrounding countryside In some cases the difference may be only nominal, for though the wages of the carpenter in the village may be lower than those obtainable for the same grade of skill in the city, his living expenses may be lower also, so that his real wages — the command over commodities which his money wages give him — are as high as those of the city laborer. But always there remain real differences which are not thus to be explained away. Labor as a matter of fact is a relatively immobile "commodity." The personal element is always present. A laborer who has estab-

lished himself and his family in a particular environment dislikes intensely to break the old ties and form entirely new associations elsewhere, and he will do so only when he becomes convinced that the benefits derived from the new location are of a very substantial nature. Wheat will be sent back and forth from one city or one country to another as the price fluctuates; labor moves much more slowly and reluctantly.

The domestic trade of the United States. From the foregoing study of general principles we may now profitably turn to a picture of the actual commerce of the United States, which falls naturally into two parts, domestic and foreign. It is noticeable that in most discussions of trade attention is apt to be directed predominantly to the country's foreign trade, so much so that many persons have an entirely distorted notion of the relative importance of the two branches of national trade. As a matter of fact the domestic trade of the United States is much larger in volume and hence presumably more significant to the welfare of the nation than the foreign trade. Unfortunately there are no adequate statistics of domestic trade. Exports and imports on the contrary are carefully registered at the customhouses. Figures on the volume of foreign commerce are therefore available and are frequently cited in the newspapers and financial journals, the course of our foreign trade is eagerly followed by business men, and as a result popular emphasis is wrongly placed.

In the endeavor to make reasonable estimates of the volume of domestic trade many ingenious methods have been utilized, but it should be remembered that all figures of domestic trade are but estimates and at best subject to a wide margin of error. Depending on the method used, the foreign trade is variously estimated as amounting to from two to nineteen per cent of the domestic trade. In other words, taking the lowest estimate of our domestic trade, it is five times the foreign trade in volume; taking the highest figures, it amounts to fifty times the foreign trade.

The United States Bureau of Foreign and Domestic Commerce in its annual survey of foreign trade attempts to arrive at an indication of the relative importance of domestic and foreign trade by comparing the values of total exports from the United States and

the output of agriculture, mining, and the manufacturing industries, plus costs of transportation. The result obviously does not give the ratio of foreign to domestic trade, but a comparison of this ratio for successive years has some significance in indicating the trend in production for the foreign and for the domestic markets. It is interesting to note that, in the thirty years or so for which these figures are available, total exports have on the average amounted to about ten per cent of total production as defined above; in 1915 the figure reached a maximum of 15.7 per cent, and in 1933, a low point of 6.6 per cent.

Foreign trade of the United States: Volume. This is not the place for an extensive survey of the foreign commerce of the United States, but the reader will doubtless be interested in some broad knowledge of its size, its general character, and its development. The table below gives the total value of exports and imports for a number of years and will serve to illustrate the phenomenal growth of this country's foreign commerce as a whole.

FOREIGN COMMERCE OF THE UNITED STATES
(in millions of dollars)

	Exports (domestic and foreign)	*General imports*	*Excess of exports (+) or imports (−)*
1800	70.9	91.0	−20.1
1820	69.7	74.4	−4.7
1840	123.7	98.2	+25.5
1860	333.6	353.6	−20.0
1880	835.6	667.9	+167.7
1900	1,394.5	849.9	+863.0
1910	1,744.9	1,556.9	+188.0
1920	8,228.0	5,278.5	+2,949.5
1925	4,910.0	4,227.0	+683.0
1928	5,129.0	4,091.0	+1,037.0
1930	3,843.0	3,061.0	+782.0
1932	1,611.0	1,323.0	+288.0
1934	2,133.0	1,655.0	+478.0
1937	3,345.0	3,084.0	+261.0

This table is interesting in showing the tremendous increase in exports and imports during the period as a whole, together with the very material decrease in foreign trade during the recent years of depression. The reader will however avoid drawing hasty conclusions with respect to the growth of trade from figures which give

values only, for there have been marked changes in price levels during the period. In general it is quite possible for the physical volume of trade to decline from one year to another, while the value of the trade increases. Professor Fisher gives the following interesting example of this:[1] "The money value of car wheels exported from the United States in one month was $12,000 and in a later month $15,000, from which fact we might infer that the quantity of these exports had increased. But the *number* of car wheels exported in the first of these two months was 2,200, and in the second only 2,100, showing a decrease. The price had increased faster than the number had decreased."

The Bureau of Foreign and Domestic Commerce has for a number of years prepared index numbers showing the changes in quantity, average unit values, and total value of the exports and imports of the United States. The reader will find it interesting to compare the changes in these index numbers, given for certain years in the table below, with the value figures given in the preceding table.

INDEXES OF CHANGES IN QUANTITY AVERAGE UNIT VALUES, AND TOTAL VALUE OF GENERAL IMPORTS AND EXPORTS OF UNITED STATES MERCHANDISE

(1923–25 average = 100)

Year	*Exports of United States merchandise*			*General imports*		
	QUANTITY INDEX	UNIT VALUE INDEX	VALUE INDEX	QUANTITY INDEX	UNIT VALUE INDEX	VALUE INDEX
1913	84	65	55	66	70	46
1920	116	156	181	88	155	136
1925	107	100	108	101	105	109
1928	128	88	113	115	92	106
1930	109	78	85	111	71	79
1932	69	51	35	79	43	34
1934	74	63	47	86	50	43
1936	82	66	54	118	51	63
1938	105	65	68	94	54	51

[1] *Elementary Principles of Economics*, p 19

[2] For a description of the methods by which these index numbers are derived the reader is referred to p 27 of *Foreign Trade of the United States, Calendar Year 1938*, Trade Promotion Series No. 198, Bureau of Foreign and Domestic Commerce, Washington, 1939

Commodity character of foreign trade. The following table gives for selected years the commodity character of the foreign trade of the United States.

COMMODITY CHARACTER OF EXPORTS AND IMPORTS IN PERCENTAGES OF TOTAL EXPORTS AND IMPORTS OF THE UNITED STATES

Commodity	*Imports*		*Domestic exports*	
	*1860**	*1937†*	*1860§*	*1937‖*
Crude materials	11%	32%	68%	22%
Crude foodstuffs	13%	14%	4%	3%
Foodstuffs wholly or partly manufactured	17%	15%	12%	5%
Manufactures for further use in manufacturing	10%	21%	4%	21%
Manufactures ready for consumption	49%	18%	11%	49%
Total (in millions)	$354	$3,012	$316	$3,295

* General imports for year ending June 30
† Imports for consumption for calendar year
§ Year ending June 30
Calendar year

These figures tell the story of a revolutionary change in the character of the foreign trade of the United States, indicating the growth in relative importance of the nation's manufacturing industries. In 1860 the nation exported 72 per cent of its goods in crude form. In 1937 all but 25 per cent of the domestic exports were partly or wholly manufactured, and the percentage of consumable manufactures had increased from 11 to 49. On the import side the converse is seen to hold true. The percentage of manufactures ready for consumption which were imported declined from 49 in 1860 to 18 in 1937, the percentage of goods partly fabricated but requiring further processes increased, and the percentage of raw materials imported increased greatly.

The direction of foreign trade. The following table shows the destination of exports and the sources of imports in recent years by great trading areas. The quantities are given both in terms of value and in percentages of the totals of exports and imports respectively, so that the reader may appreciate both the absolute and the relative importance of the different regions of the world in American foreign trade.

We cannot attempt an exhaustive study of the trade relationships revealed by this table, but attention should be called first of all to the great, though declining, importance of Europe in American

FOREIGN TRADE BY GREAT TRADE REGIONS
(Annual averages. Values in millions of dollars)

Year	*North America* Northern Value	Northern Per Cent	Southern Value	Southern Per Cent	*South America* Value	Per Cent	*Europe* Value	Per Cent	*Asia* Value	Per Cent	*Oceania* Value	Per Cent	*Africa* Value	Per Cent
Exports														
1910–14	320	14.8	181	8.4	121	5.6	1,350	62.3	121	5.6	48	2.2	25	1.1
1921–25	627	14.3	445	10.1	297	6.8	2,318	52.7	499	11.3	141	3.2	70	1.6
1926–30	830	17.4	403	8.4	448	9.4	2,236	46.8	574	12.0	177	3.7	110	2.3
1932	246	15.3	119	7.4	97	6.0	784	48.7	292	18.2	37	2.3	36	2.2
1934	308	14.4	178	8.3	162	7.6	950	41.5	401	18.8	57	2.7	77	3.6
1936	392	15.9	225	9.2	204	8.3	1,043	42.5	399	16.2	79	3.2	114	4.7
Imports														
1910–14	119	7.0	229	13.5	207	12.2	836	49.5	259	15.3	17	1.0	23	1.3
1921–25	397	11.5	514	14.9	421	12.2	1,049	30.4	943	27.3	54	1.6	71	2.1
1926–30	480	11.9	460	11.4	546	13.5	1,211	30.0	1,193	29.6	53	1.3	91	2.3
1932	181	13.7	157	11.9	201	15.2	389	29.4	362	27.4	8	0.6	24	1.8
1934	238	14.4	161	9.7	229	13.8	489	29.6	490	29.6	15	0.9	33	2.0
1936	381	15.7	237	9.8	292	12.0	718	29.6	708	29.2	36	1.5	51	2.1

foreign trade. In 1936 she took about 43 per cent of the total exports of the United States and provided nearly 30 per cent of the imports. Of the total exports to Europe 40 per cent was in the form of crude materials, foodstuffs amounted to 15 per cent; semi-manufactured goods, to 16 per cent; and finished manufactures, to 29 per cent. As might be expected, imports from Europe consisted chiefly of manufactured goods, both finished and partially finished.

The importance of the trade with North America, attributable almost entirely to the trade with Canada, is striking. Commerce with the southern part of North America — and the Cuban trade bulks largest in this — shows a decline, both absolutely and relatively, after a period of slight growth. The trade with the countries

of South America, on the other hand, has grown materially since 1910. In part this may be ascribed to the privileged position which the United States enjoyed during the World War, but in the main it is a reflection of the complementary nature of the production of the United States and of many countries of South America. A very natural basis exists for the exchange of the coffee, nitrates, tin, hides, etc., of South America for the manufactured goods of the United States.

Another feature deserving of mention is the position of Asia. In the twenty-seven years covered in the table, American imports from Asia have more than doubled, and exports to Asia have about tripled. Raw silk, rubber, tin, furs, hides, and skins, jute, and other raw materials predominate in the imports from Asiatic countries, while exports to this continent consist mainly of manufactured goods.

Imports. The table below shows the value of the ten leading articles imported into the United States in 1936.

LEADING IMPORTS OF THE UNITED STATES IN 1936
(in millions of dollars)

Crude rubber	158.7
Cane sugar	157.9
Coffee	134.0
Paper and paper manufactures	110.1
Raw silk	102.4
Paper base stocks	98.9
Vegetable oils	85.3
Furs and fur manufactures	81.6
Chemicals and related products	80.0
Tin	75.5
Total	1,081.4
Total imports	2,423.0

Judged on the basis of value these ten items are the most important articles imported. A mere glance at the list will suggest how essential most of these articles are to the economic life of the people. But it would require a painstaking study of the hundreds of different items coming from all parts of the world to enable one to appreciate fully the extent to which economic life is dependent on foreign sources of supply. In addition to the goods already mentioned, we would find many minerals and other raw materials

which are produced not at all or in insufficient quantities in America, such as nickel, manganese, antimony, jute, and flax. There would be many manufactured goods which perhaps the American people could produce but which they can get with less effort by trade, such as fine woollens from England and watches and lace from Switzerland. We would discover also many articles of food — tea, coffee, cocoa, and out-of-season fruits and vegetables — which could be secured at home only at prohibitive cost.

Of recent years there has been much talk of the advantages of a self-contained nationalistic economy, to be attained presumably by the erection of prohibitive tariff barriers. A study of the problems raised by the protective tariff will be taken up in a succeeding chapter, but it may not be out of place here to point out that a study of its imports soon convinces one that even the United States, with its wide expanse of territory and its diverse natural resources, could not hope to proceed very far in that direction unless the people were willing to alter substantially their mode of life and to submit to a distinctly lower standard of living.

Exports. The following table gives the values of the ten leading exports from the United States in 1936.

LEADING EXPORTS OF THE UNITED STATES IN 1936

(in millions of dollars)

Cotton, raw	361 0
Machinery	334.9
Automobiles	240 2
Petroleum and its products	179 0
Tobacco, unmanufactured	137.3
Chemicals and related products	116 9
Iron- and steel-mill products	111 9
Fruit and nuts	80.6
Coal and coke	56.6
Copper, ore, and manufactures	50 5
Total	1,668 9
Total exports	2,456 0

The sum total of the exports is a part of the price which has to be paid for the imports and other services from abroad.[1] They are the goods in the production of which American advantages are greatest. By devoting time and energy to their production the

[1] Other items than exports and imports which enter into international financial relationships are treated in Chapter XXXV.

imported goods are secured on easier terms than would be the case if they were to be produced at home (assuming this to be possible) or if capital and labor were to be devoted to making other goods in America for sale abroad.

It is because a country's exports represent the most efficient means of acquiring certain desired goods that every country is properly concerned with the condition of its export trade. Any factor, domestic or foreign, which interferes with the free outflow of exports decreases the purchasing power of a nation in foreign markets and is a legitimate cause of concern.

There is another reason why a nation may view the volume and character of its export trade with some anxiety. While it is true that production for the foreign market is on the average a fairly small part of total production, in particular industries it may amount to a very large fraction of total output. The following table gives for selected industries the percentage of total production in the United States which was exported in the years 1914, 1929, 1933, and 1936.

UNITED STATES EXPORTS OF LEADING COMMODITIES IN RELATION TO PRODUCTION

	Percentage exported			
	1914	*1929*	*1933*	*1936*
Tobacco, leaf*	47.2	41.2	39.3	33.4
Cotton†	62.6	54.8	65.6	56.1
Wheat‡	19.7	17.9	5.4	2.5
Copper, refined	54.8	36.2	41.0	32.1
Refined mineral oils	37.9	19.6	11.9	9.9 §§
Industrial machinery	--	13.3	11.2	13.3 §§
Agricultural implements and machinery	—	25.1	29.8 §	10.6 §§
Automobiles	4.5	12.0	7.3	8.0
Office appliances	23.6	30.2	26.1	23.2 §§

* Based on production figures for the preceding calendar year and exports for the year ended June 30

† Based on production figures and exports for the year ended July 31.

‡ Exports include flour

§ Data are for 1931

§§ Data are for 193[illegible].

These are the more conspicuous examples, but many industries could be cited which have customarily exported up to ten per cent or so of their total output. Quite naturally anything which interferes with the foreign market of such industries is a serious matter,

for the loss of ten per cent, or even less, of the total market may mean the difference between producing at a profit or at a loss. The loss of foreign markets is of immediate practical concern, not alone to the particular producers immediately involved, but to the nation as a whole. If foreign markets are closed to an appreciable extent, the depressing influence is not confined to the export industries but tends to be generalized and cumulative. In the course of time labor and capital, thrown out of use in the export industries, can be diverted to other uses and equilibrium achieved anew, but not without irreparable loss.

XXXIV

FOREIGN EXCHANGE

International payments. Payments between persons in the same place are made with money or checks, and the process has been sufficiently investigated. Payments between persons in different places, on the other hand, involve certain special problems, because of the necessity of making currency paid in one place available in another place. If in addition the two persons are in different countries, a further complexity may arise because of the difference in monetary units. Since all the economic problems involved in payments between places in the same country are included in the more complex subject of international payments the whole matter may properly be discussed under the subject of foreign exchange.

In our first approach to the study of foreign exchange it will be convenient to eliminate for the time being certain complicating factors of recent years. The difficulties which confront the student of contemporary foreign exchange arise partly from the abandonment of the gold standard by most of the countries of the world and partly because of a variety of controls, exercised by governments or by central banks in pursuance of governmental policy, affecting exchange operations or transactions which would involve the exchanges. We shall consider these problems in a later chapter, but for the present we will assume (1) that the countries with which we are concerned are on the gold standard, implying thereby that all forms of money within each country are ultimately redeemable in the country's standard gold coins and that the right to possess, export, and import gold may be freely exercised by private persons, and (2) that there are no formal exchange controls which interfere with the operations of a competitive market in foreign exchange. Such were the conditions which existed with respect to the exchanges of all the leading industrial nations in

the pre-war period, and it is quite possible that they may again be attained.

Financing exports. Our first problem — one which is much the same whatever the monetary systems on which the exchanges are based — is to investigate the mechanism which is available for financing exports and imports of goods. The simplest approach to this is by way of the study of typical examples of export and import transactions. Let us therefore assume that Richards & Green of New York has sold Harris & Company, Ltd., of London a certain number of bales of cotton for £1,000. Now it is evident that the ordinary forms of currency used in domestic trade are not available for payment in this case. It is within the realm of possibility that Harris & Company, Ltd., might effect payment by shipping a quantity of gold to Richards & Green. Gold is a commodity which has a ready market, and were Richards & Green to receive one thousand gold sovereigns it could sell them to bullion brokers or take them to the mint and have them coined into American dollars. But the shipment of gold is a troublesome and expensive operation, which is practically never used for commercial payments. The universal currency for international payments is the bill of exchange.

We may assume — what would almost always be true in such a case — that the terms of the sale have called for a bill of exchange to be drawn by Richards & Green on Harris & Company, Ltd. The bill will be in the following form:

> No. 726 New York, December 29, 1929
>
> Pay at sight to the order of the New York Bank
> One Thousand Pounds Sterling
>
> (signed) Richards & Green
>
> To Harris & Company, Ltd
> London, England

The bill of exchange is an instrument of such importance that we must pause here to formulate its precise technical definition, which is as follows: *A bill of exchange is an unconditional order in writing,*

signed by one person (called the drawer) directing a second person (called the drawee) to pay a third person (called the payee) or his order a certain sum of money at some definite time in the future. It will be noted that the example given above conforms with this definition. The bill of exchange appears in various other forms, of which the simplest is the ordinary bank check, with which the reader has become acquainted through the medium of an earlier chapter, and doubtless also from practical experience.

When Richards & Green ships the cotton it secures from the steamship company, or possibly from a railroad company if the shipment is made from an inland warehouse, a bill of lading, which not only shows that the company has received a certain number of bales of cotton and has agreed to transport them to England, but also serves as the evidence of ownership in the cotton. Anyone in possession of a properly endorsed bill of lading can secure the cotton from the steamship company, and it can be secured in no other way. Marine insurance must be taken out on the shipment in order that the owner may be protected if the cotton suffers loss or damage during the sea voyage, for ocean carriers are by law exempt from practically all responsibility with regard to the safety of shipments entrusted to them.

Richards & Green next takes the bill of exchange, with these two documents (bill of lading and insurance policy) and possibly others attached, to its bank, the New York Bank, for sale. The bank knows nothing about the reliability of Harris & Company, Ltd., but it knows that Richards & Green is an old, well-established house with a reputation for financial straightforwardness. If it buys the bill of exchange it will acquire a credit instrument bearing the signature of a responsible officer in the concern and having the same standing as a promissory note signed by the same officer. If Harris & Company, Ltd., fails to meet the bill when presented for payment the bank can enforce payment from Richards & Green. Furthermore the bank acquires possession of the bill of lading, which will give it the title to the cotton. Now cotton is a staple article, not perishable, and commanding at once a definite price in all the important markets of the world. If the worst should happen and the bank could not collect from Richards & Green,

it could sell the cotton and still not lose much. Relying therefore on the responsibility of Richards & Green and on the cotton as collateral security, it purchases the bill of exchange at the rate prevailing in the market at the moment for bills of that character. The rate, let us say, happens to be \$4.84 [1] per pound sterling, and Richards & Green therefore receives \$4,840 in cash. The exporting firm has thus received payment for the cotton in American dollars immediately after shipment and has no further interest in the transaction, save for its contingent liability.

The bank now sends the bill, with the documents still attached, to the London Bank, with which it has correspondent relations, with instructions that it be presented at once to Harris & Company, Ltd., for payment and that the proceeds be credited to the account of the New York Bank. Harris & Company, Ltd., pays the bill when presented and receives the bill of lading, which entitles it to claim the cotton from the agents of the steamship company.

We may summarize these stages in the financing of the export of cotton as follows: (1) Richards & Green draws a bill of exchange on Harris & Company, Ltd., for £1,000. (2) Richards & Green sells the bill of exchange (with documents attached) to the New York Bank for \$4,840. (3) The New York Bank sends the bill and the documents to its correspondent in London, the London Bank. (4) The London Bank collects £1,000 from Harris & Company, Ltd., and surrenders the bill of lading. (5) The New York Bank's deposit account with the London Bank is credited by £1,000.

Financing imports. A similar example will aid us to visualize the several steps involved in the financing of an import. We assume that White & Company of New York has purchased £800 worth of cotton goods from R. J. Littlefield of London. It is again possible for White & Company to ship a sufficient quantity of gold to be worth £800 in London, or Littlefield might draw a bill of exchange on White & Company, thus reversing the process described in the previous section. But as a matter of fact trade

[1] The student will realize that this rate reflects conditions of 1929 when both Great Britain and the United States were on the gold standard.

practice frequently places the responsibility of making payment, as well as of securing payment, on the American merchant. So in this instance White & Company is required to place Littlefield in possession of £800 in cash or its equivalent in London. This it can do by purchasing from its banker a bill of exchange such as the following, often called a demand or sight draft.

No 918 New York, January 5, 1930

At sight pay to the order of R J Littlefield
Eight Hundred Pounds Sterling
and charge the same to our account

(signed) New York Bank

To
The London Bank

The New York Bank sells this bill of exchange to White & Company at the prevailing rate for bankers' demand drafts, which we will assume to be $4 87. The draft therefore costs White & Company $3,896. White & Company sends the draft by mail to Littlefield, who presents it at the London Bank and receives £800 in cash, which amount is debited to the account of the New York Bank.

Various types of bills of exchange. In the examples of bills of exchange only one type has been mentioned, the demand draft, of which there are two forms, the commercial demand draft, drawn by one commercial house on another, and the bankers' sight draft, drawn by one banker on another. It is not necessary to attempt a classification of the varied forms which the bill of exchange in foreign trade may take, beyond noting the fact that such forms exist and are in common usage wherever the foreign bill of exchange is used. Attention should however be called to a very common form, the time or long bill of exchange, which requires payment a certain number of days after date, or after presentment or *sight*, as it is called. If Richards & Green had sold its cotton to be paid for on the basis of a time bill of exchange, the bill might have read as follows:

No. 701 New York, Dec. 29, 1929

Sixty days after sight pay to the order of
The New York Bank
. . . . .One Thousand Pounds Sterling . . . .

(signed) Richards & Green

To Harris & Company, Ltd.
London, Eng.

The New York Bank will be unwilling to pay the same rate for this bill of exchange that it paid for the sight draft, because the bill will not be payable until sixty days after it has been presented to Harris & Company, Ltd., for acceptance. To determine its value at the moment of acceptance one must discount it for sixty days at the rate of discount prevailing in the London discount market. The rate which the New York Bank pays will take this fact into account in a very precise manner and will be lower than the sight rate, say $4.81 per pound sterling. When this bank sends the bill to the London Bank, it gives instructions to present it to Harris & Company, Ltd., for *acceptance*, which means that Harris & Company, Ltd., will write on the face of it "accepted and payable at the London Bank on such and such a date" and sign it, thereby formally acknowledging its obligation. The New York Bank may then request the London Bank to have the bill discounted and the proceeds placed to its credit, so that it may draw its own drafts against the credit, or it may conceivably desire to have it held until maturity, thereby earning the interest for the two months. Probably the largest portion of all American exports is paid for by time bills of exchange of one sort or another. They are therefore of considerable importance, but it should be remembered that they differ from the sight draft only in the time element, and that they introduce no new problems in the theory of foreign exchange.

The rôle of the banks in exchange. The part which the bankers play is that of middlemen. It would indeed be possible for importers to buy bills of exchange from exporters and to that extent eliminate the bankers, and in the past this has been done. We

know that in the first decade of the nineteenth century American importers rarely paid by bankers' drafts; in almost all cases the bills remitted were those drawn on British purchasers by exporters of cotton, grain, or other products. Today virtually all bills are in the first instance either sold to or bought from the bankers. The banker has a wide circle of customers importing and exporting every known ware and dealing with all the markets of the world, and he is able to buy and sell drafts in all currencies and in all amounts at a moment's notice. It is patently much simpler for an importer who must remit a draft for £168 9*s.* 3*d.* to purchase it from a banker than to find an exporter who has a draft for this precise amount, for the banker can draw a draft in any denomination so long as he has a credit balance abroad sufficiently large to cover it, while the size of the exporter's draft is regulated by his export. An importer of goods from China can secure a draft payable in Shanghai or Hongkong with no difficulty, and furthermore he knows and the Chinese exporter knows that the risk of loss inherent in a banker's draft is much smaller than in a draft drawn by one commercial house on another. Because of these advantages it is not surprising to find the banker at present the focal point of the foreign exchange market. Only rarely would an importer who is required to remit a bill of exchange send anything but a banker's draft, or an exporter who has a bill for sale look for a purchaser outside his bank.

Middlemen always expect to sell what they buy, and when bankers buy bills of exchange it is because they expect to be able to sell bills of exchange. We have seen that the result of the purchase by the New York Bank of the bill drawn by Richards & Green was to give this bank a credit with the London Bank of £1,000. It was the existence of a credit such as this which enabled the New York Bank to sell its own demand draft to White & Company. By the same token the sale of this draft to White & Company gave the bank cash wherewith to purchase other bills of other exporters. Every time a bank buys a bill payable in London, it reduces the supply of cash in its own vaults and acquires a deposit in England. It will not buy bills of exchange simply to pile up unused credits in the London Bank, for the low rate of interest which it might get

on such a deposit would not compensate it for the loss of funds in New York, and furthermore so one-sided a business would soon drain all its cash resources and compel it to suspend operations. On the other hand, every time a bank sells a bill payable in London, it acquires cash in its own vaults and reduces its deposits in England. The bank cannot long continue a one-sided business of this sort without entirely exhausting its London credit and so being compelled to suspend operations.

If however the bank regulates its purchases of bills and its sales of its own drafts in harmony with each other, it will create and utilize its foreign credits and expend and replenish its cash resources at home in approximately equal amounts from day to day. These balanced transactions will permit it to keep its balance with its foreign correspondent neither unduly large nor unduly small and likewise to avoid both a shortage of cash and a surplus of idle funds in its home office. The bank dealing in foreign exchange intends that its operations shall approximate this equilibrium, while making its profit in foreign exchange by selling at a slightly higher rate than the one at which it buys.

The exchange market. The exchange rate responds in so sensitive a manner to the play of the forces of demand and supply and there is such a vast quantity of transactions in foreign exchange in New York and other centres, that it is frequently assumed that there must be an organized market similar to the stock and produce exchanges where the rate of exchange is determined. This is an error; there is in New York no formal organized exchange for dealing in foreign exchange. The bulk of the business is concentrated in the hands of a relatively few large banks and exchange houses. Each bank fixes the rates at which it will buy and sell exchange on all the financial centres with which it has dealings, but each house keeps so closely in touch with the activities of the other houses that there is seldom any material discrepancy between the rates charged by the various dealers on the same class of bills.

There is no official rate of exchange. Each bank posts the rates at which it is willing to buy and sell bills of exchange, but it frequently departs from its posted rates both on bills of large amounts and unquestioned character and on bills of dubious standing. In

fact some bills will not be bought at all, but accepted for collection only. Since 1921 the Federal Reserve Bank of New York has been required to issue a daily report of the buying rate in New York for cable transfers on the important financial centres, and this possibly is the nearest approach to an official quotation. But it should be borne in mind that this rate is not a rate at which the Federal Reserve Bank itself is offering to buy cable transfers but is one obtained from the regular dealers in exchange.

The rate of exchange: General principles. Thus far we have been concerned chiefly with the technical facilities which modern business uses in the transmission of funds from one country to another, but we are equally interested in the forces which determine the *rate of exchange*. This may be defined as *the price paid in the home country for one unit of the money of the foreign country payable in that country*. For example the rate of exchange in New York on London is the amount of American money that must be paid in New York for a bill of exchange calling for the payment of one pound sterling in London.[1]

The first step in the study of the rate of exchange is to determine what is known as the *mint par of exchange*. This is a simple ratio derived by dividing the fine gold content of the standard unit of one country by the fine gold content of the unit of another. Italy and Rumania, for example, used to have gold coins containing the same number of grains of fine gold, the mint par of exchange between them was therefore one; *i.e.*, one lira was equivalent to one leu. Formerly the sovereign of Great Britain contained $113.00116\frac{2}{3}$ grains of fine gold, and the gold dollar of the United States contained 23.22 grains of fine gold. The ratio was $\frac{113.00116\frac{2}{3}}{23.22}$ or 4.8665 +. Hence the British sovereign contained 4.8665 + times as many grains of fine gold as the dollar, and the mint par of exchange was $4.8665 + per pound sterling. This means that if ten thousand gold sovereigns of full weight were to

[1] Rates of exchange may equally well be expressed the other way round, as the amount of the foreign money payable in the foreign country which may be obtained in the home country for one unit of its money. Thus before the World War the rate of exchange on France was often stated as so many francs per dollar. This is not the usual method however.

have been taken to the United States mint they would have contained sufficient fine gold to have furnished the gold content of $48,665 +.

The rate of exchange however is not the same thing as the mint par of exchange. Occasionally the rate of exchange quoted in the market stands at the mint par, and when it does so we say that exchange is "at par." But this rarely occurs, and then only by chance. The mint par of exchange furthermore does not determine the rate of exchange; its significance lies, as will appear, in the fact that it helps to set limits to the fluctuations of the actual rates.

The rate of exchange is a price and, in common with competitive prices in general, is determined by demand and supply. The demand for bills of exchange comes primarily from the importers and others who are under the necessity of making payments in foreign countries. The supply comes primarily from the exporters and others who have money available to them in foreign countries. The bankers are the middlemen who, buying and selling, bring together the forces of demand and supply. In accordance with the general laws of price, the greater the demand for bills (other things being equal), the higher the rate of exchange; the greater the supply (other things being equal), the lower the rate. In terms of international trade, when imports are large compared with exports, the rate of exchange tends to be high; when exports are relatively large, the rate tends to be low.

The gold points. But the rate of exchange differs from most competitive prices in one important respect; between countries on the gold standard it can fluctuate only within definite limits. Having risen to the upper limit, the rate will stand there, no matter how great the demand, until changed conditions send it down. Having fallen to the lower limit, it can go no lower and will eventually rise. To make this clear let us assume that, because of the transactions of a certain short period, business men in America owe £50,000,000 on account of imports received and are owed only £25,000,000 on account of exports. By buying the bills of the exporters, American bankers can acquire credits in London of £25,000,000 and be in position to sell their own bills to this amount to the importers. But this will leave £25,000,000 of the debts

of the American importers unpaid. As this discrepancy between the volume of bills which importers and others want and the volume offered to the bankers for sale begins to be manifest, bankers will tend to charge a higher price for their own bills of exchange in order to protect their foreign balances; and individual bankers, each one of whom will want to get as large a share of the possible credits abroad as he can, will bid against each other for the bills of exchange which are offered for sale by exporters. The rate of exchange will be forced up gradually, and the higher rate will tend to discourage some who want to make payments and perhaps encourage others to anticipate the drawing of bills of exchange. But if the basic forces at work are strong enough the bankers will soon find their foreign balances exhausted, with importers still trying to buy bills of exchange. In this situation the bankers in America, if they are not to cease operations entirely, are forced to resort to other expedients to build up their foreign bank balances. They might acquire balances against which they could draw by borrowing from a bank in England or by exporting securities for which there is an international market, but, granted our assumption of a free gold market, they would in all probability export gold from the United States for this purpose.

At a mint par of exchange of $4.866 per pound sterling, gold equivalent to one thousand pounds sterling can be purchased in the United States for $4,866. It will obviously cost something to cover the shipment of gold — to take care of crating, freight, insurance, and the like. Assuming that these expenses amount to two and one half cents per pound sterling, then the total cost of placing one thousand pounds in London would be $4,891. The bankers could therefore offer to sell bills of exchange drawn on banks in London at $4.891 per pound, and the importers who were under obligations to make immediate payment would have to pay this rate. True they could ship the gold themselves, but the banks can do this much more cheaply, and competition among different banks forces them to do so. This rate we call the *gold exporting point,* determined by the cost in New York of transferring one gold sovereign to London. It is the highest price which will be paid in one country for one unit of a foreign currency payable in the

foreign country and, no matter how great the demand, bankers will continue to furnish bills of exchange to all comers at that rate so long as gold is available for shipment.

In like manner we may determine the *gold importing point.* Let us suppose that the supply of sterling bills of exchange; that is, bills drawn on London in pounds sterling, increases greatly, without any corresponding change in the demand. As we have seen, the supply of bills of exchange arises when exporters and others having debts to collect draw bills of exchange against the foreign debtors and attempt to sell them to bankers, and a demand arises when importers and others having debts to pay to foreigners go to the bankers to buy bankers' drafts. In this case therefore bankers are likely to have large credit balances abroad for which there is no call, and if the situation continues they will refuse to buy bills except at rates which will permit them to transfer their foreign balances to replenish their dwindling domestic cash resources. The cost of shipping gold from England to America is slightly higher than in the other case, for in this instance the bankers lose interest during the time of shipment; whereas, when gold is shipped from New York, the bankers can sell drafts to go by the same steamer and so avoid tying up their funds. If we assume the cost of shipment to be three cents per pound sterling, bankers will be willing to buy bills of exchange at the rate of $4.836. With a draft of £1,000 a banker can obtain a thousand gold sovereigns, bring them to the United States, and sell them for $4,866. The cost of doing so being $30, he will recover $4,836, the amount he paid for the bill. This rate is known as the gold importing point, and at this rate bankers will be willing to purchase an unlimited quantity of bills of exchange, no matter how great the supply.

The gold points then are determined with respect to the mint par of exchange and fix the limits within which the rate of exchange may fluctuate. If the rate rises to the upper gold point, gold will be exported; if it falls to the lower gold point, gold will be imported The rise or fall in the rate is determined by the relation of demand and supply to each other, which in turn rests on the relation between our obligations to foreigners and the obligations of foreigners to us. If these were just equal, the rate would presumably

be at or very near the mint par of exchange, and there would be no tendency for gold to flow at all; if at any time one greatly outbalances the other, gold will flow out to create the necessary foreign bank balances, or gold will flow in as bankers find this the only way to utilize their foreign balances.

Indirect exchange. We have thus far examined the machinery by which the obligations arising out of the trade between two countries are settled, as though each country had to make a separate settlement with each other country, much as the banks made separate settlement with each other before the days of the clearing house. If this were the actual case, there would be almost continual shipping in and out of large quantities of gold, since it is scarcely conceivable that the import and export transactions between any two countries should exactly offset each other except as a rare coincidence or, as in certain recent cases, by official interference or control. We should expect to see the United States, for example, receiving large shipments of gold from Great Britain while at the same time it was making heavy shipments to Brazil. As a matter of fact, when the majority of the countries of the world were on the gold standard, international trade was always conducted on a clearing house basis, in accordance with which all the obligations of the residents of one country to residents of all other countries were offset against all claims against foreigners. There was no international clearing house, but the result was accomplished with remarkable certainty and smoothness through the foreign exchange markets of the world. To a great extent this is still true, although as we shall see in a succeeding chapter, many barriers have been erected which impede the smooth operation of the system.

In this system London and sterling exchange played an important rôle. The leading position of Great Britain in world trade, the solidity and flexibility of her monetary and banking system, and the supremacy of London in international short-term and long-term lending as well as in other financial services gave to London a dominating position in international finance, which only recently she has had to share with New York and Paris. These circumstances gave to sterling exchange the character of an international

currency. For many years bills on London were in all markets a most acceptable form of exchange. Indeed so firmly was this tradition established that in the settlement of many American obligations sterling bills of exchange were preferred to dollar bills or even to bills drawn in the currency of the local government.

An illustration of the workings of indirect exchange, utilizing sterling exchange as an international currency, can be seen in the trade between Great Britain, Brazil, and the United States. The American people normally import more from Brazil than they sell to her, they export more to Great Britain than they purchase from her, and the people of Great Britain sell more to the Brazilians than they purchase from them. If each country were to settle with each separately and without regard to its financial obligations to other countries, the United States would be required to send a large sum in gold annually to Brazil, Brazil in turn would have to settle in gold with Great Britain, and a third shipment of gold would be necessary from Great Britain to the United States. Or if gold could not be shipped, Americans would cut down their imports from Brazil until they balanced the exports to her, Brazil in turn would decrease her imports from Great Britain until they were in balance with her exports to Great Britain, and American exports to Great Britain would likewise be decreased. The second alternative obviously would mean a tremendous decrease in international trade, with unfortunate reactions on international division of labor. The obligations created by this three-cornered trade can however be settled much more simply. America's excess of exports to Great Britain gives bankers in the United States balances in London banks which can be utilized in drawing bills of exchange in favor of Brazilian exporters. In due course these drafts will be sold by the Brazilian exporters to bankers in Brazil and sent to London for collection. Brazilian bankers therefore have acquired balances in London banks against which drafts can be sold to Brazilian importers of goods from Great Britain for transmission to the British exporters. It is possible, though very unlikely, that the surplus balances acquired from American exports to Great Britain might be exactly the amount necessary to pay for the excess of imports from Brazil; it is more likely that Americans

would still have a balance in London, which could be utilized to pay for an excess of imports from Colombia, let us say.

The existence of an exchange such as sterling exchange which is everywhere preferred is not an indispensable prerequisite to international clearing arrangements of the character illustrated above; American importers could remit drafts expressed in terms of dollars or in terms of Brazilian currency and drawn on the London banks if these were preferred in Brazil. The existence of a stable and generally acceptable form of exchange does however make for smoothness in international financing.

EXERCISES

1. Mr Jones, an exporter of grain from the United States to England, sells a shipment worth $10,000 to Mr Blackwell, agreeing to receive payment by drawing on the buyer. Draw the bill of exchange, and trace the steps by which the transaction is handled.

2. Suppose that a New York City bank will pay $4.80 per pound sterling for a demand draft drawn by an American exporter on an English importer.
   (a) What price would the bank pay for a time draft payable 90 days after sight, if the rate of discount were four per cent? two per cent?
   (b) How does a decline in the rate of discount affect the "spread" between the prices of these two types of bills of exchange?

3. Suppose that a New York City bank asks $3,600 for a bankers' sight draft in the amount of 10,000 rupees drawn on Calcutta. How much would the buyer have to pay for a 10,000 rupee draft payable 90 days from date? Assume that the rate of interest is four per cent, and the mail time from New York to Calcutta is 30 days.

XXXV

THE INTERNATIONAL BALANCE OF PAYMENTS

Mercantilist fallacies. For centuries popular belief has clung to the notion that an excess of merchandise exports over imports always causes gold to flow into a country. The Mercantilists were convinced that the prosperity of a country depended in large measure on the ability to pile up precious metals within the country, and they directed their domestic and foreign policy with this end in view. To them, as to their present-day counterparts, it seemed an obvious fact that if the exports of merchandise were greater than the imports of merchandise the balance must be paid in precious metals. The logical policy, it seemed, was to stimulate exports and repress imports in order that this supposed desirable end might be attained. A "favorable" balance of trade, so called — an excess of merchandise exports over imports — was and still is considered by many a sign of national prosperity; an "unfavorable" balance of trade, a token that something ought to be done.[1] There are three fallacies in the concept of a favorable balance of trade. The first is in conceiving the exports and imports of a country to be the only items in the balance sheet, the second lies in the concept of gold as the most desirable of all imports, and the third is the supposition that a steady inflow of gold can be produced by artificial means.

In this chapter we shall retain our assumption that a free gold market prevails in all of the countries concerned; *i.e.*, that each country's currency is either gold or redeemable in gold and that there are no legal obstacles in the way of the export and import of gold. As a safeguard against a common error it may not be amiss to remind the reader that when one says that America exports

[1] When reference is made to the balance of trade or the balance of payments, the terms "active" and "passive" are frequently used today in place of the terms, "favorable" and "unfavorable," which tradition has handed down.

goods to England or that Brazil owes England a certain sum, reference is not to transactions between the respective governments but to all transactions private and public between parties in one nation and parties in the other nation.

The balance of payments: Balance of trade and balance of payments. *The difference between the monetary value of a country's merchandise exports and the monetary value of its merchandise imports is called its balance of trade.* It is this balance which, according to the Mercantilist doctrine, is supposed to control the international flow of gold. The error here is evident as soon as one has a clear picture of the various items which, as a matter of fact, enter into the dealings of the people of one country with those of other countries. These international transactions naturally include the exports and imports of merchandise. But there are many items other than merchandise which must enter into the balance sheet of any nation and help to determine whether or not it will export or import gold. For example the people of the United States incur indebtedness to the people of other countries because of a variety of services performed by them, and the same holds true of indebtedness of other countries to the United States. The ships of Great Britain carry a large share of America's foreign commerce, the greater portion of marine insurance premiums must be paid to the insurance companies of Great Britain, foreign banks charge fees for handling the collection of bills of exchange in foreign markets. Americans purchase securities issued by foreign governments and corporations, and foreigners likewise invest in American securities. Payments must be made for these purchases and sales and also for the interest charges and dividends to which they give rise. American tourists require large sums to support them while they are abroad, immigrants want to transmit funds to their relatives in the home country, and there are many other transactions which in the normal course of events necessitate the transfer of funds either to or from the United States. Any of these transactions may require the use of foreign exchange and, by influencing either the demand or the supply side of the market, may influence the rate of exchange and consequently the international flow of gold.

Let us suppose that an American corporation owes interest payments to bondholders in England. It will designate some bank in

England to serve as its agent in making payments to individual bondholders and will place this bank in possession of sufficient funds to enable it to redeem all the interest coupons presented for payment. In order to remit the necessary funds, the corporation will apply to some New York banker for a draft on London, which it can remit to its London agent. In other words, the obligation of the American corporation to make interest payments in London has given rise to a demand for bills of exchange on London and will exert the same influence on the rate of exchange as an importation of goods from England. If an American underwriting syndicate were to sell a British investment house a block of bonds in a new American corporation, it would thereby acquire the right to draw bills of exchange on the investment house; it would have bills to sell to American bankers and would in this manner tend to increase the supply of bills in the American market. In precisely the same way any transaction which places any person in England in the debt of an American will give rise to a supply of bills of exchange on London. Items of this character may be brought together in the form of a country's international balance sheet, by listing as credits the transactions for which foreigners must make payments to this country and as debits those which represent obligations by residents of this country to foreigners. *The difference between the money value of all items for which residents of this country must pay foreigners and the money value of all items for which foreigners must pay this country's residents (save gold, the balancing item), is known as the balance of payments.*

Balance of payments illustrated. For the sake of a mathematical example let us assume that A and B are countries trading with each other exclusively and that the people of A in a particular year export \$100,000,000 of goods to the people of B and import goods valued at \$75,000,000. We find also that people in B have rendered for those of A shipping services worth \$15,000,000, that the bondholders in B are owed \$25,000,000 in interest, and that emigrants from B have made remittances of \$5,000,000. On the other hand, the people of A by carrying freight have performed services worth \$5,000,000, they are owed \$5,000,000 in interest on the foreign bonds which they hold, and emigrants from A have remitted

$10,000,000. The balance sheet of A will be as shown in the paragraph below.

The bankers in A can secure credits of $120,000,000 in B by the purchase of the bills of exchange offered for sale by the exporters of merchandise, the carriers of freight, and the others who have the right to draw bills on foreigners; and the bankers can also sell $120,000,000 of bills of exchange to the importers, the corporations,

INTERNATIONAL BALANCE SHEET, COUNTRY A

Credits	
Exports of merchandise	$100,000,000
Ocean freight	5,000,000
Interest	5,000,000
Emigrant remittances	10,000,000
	$120,000,000
Debits	
Imports of merchandise	$ 75,000,000
Ocean freight	15,000,000
Interest	25,000,000
Immigrant remittances	5,000,000
	$120,000,000

the shippers, and the immigrants who have foreign payments to make. With this equilibrium of debits and credits, there is no occasion for the rate of exchange to be either much above or much below par, and since gold will not be imported until the rate falls to the gold importing point or exported till the rate rises to the gold exporting point, there is nothing to occasion a flow of gold either way.

The international balance sheet of the United States. The picture of the international financial relations of a great nation can be made clearer by a glance at the table on page 139, derived from figures compiled by the United States Department of Commerce, showing the international balance sheet of the United States for the year 1928.[1]

In this table, as in the illustration above, every transaction which creates in the United States a claim against some foreigner is con-

[1] The balance sheet for 1928, rather than one of a later year, is considered more useful for purposes of illustration since it antedates the months in 1929 when speculation was at its height, the years of depression, and the abandonment of the gold standard by various nations

sidered a credit. Such items furnish the supply of bills of exchange on foreign countries, or a demand of foreigners for bills on the United States. Thus the item "Earnings of long-term investments" means that individuals or corporations in the United States acquired credits abroad amounting to $817,000,000 because of interest and dividends from their holdings of foreign securities. Similarly every transaction which gives rise to a claim of some foreigner against an American is placed in the debit column of our account. Thus shipping services performed by foreigners cost $227,000,000, and the cashing of the drafts to which the payment of this amount gave rise decreased the foreign bank balances of Americans by this amount. Such items furnish the demand for foreign bills of exchange, or the supply of bills on the United States available to foreigners.

International gold movements: Balance of payments. The reader will have noted in the example given on page 137 (1) that exports of merchandise are larger than imports of merchandise, (2) that credits and debits exactly balance, and (3) that there has been no transfer of gold. In this instance there was a favorable balance of trade, but no importation of gold; obviously there was no need for a shipment of gold either to provide bankers with foreign bank balances or to enable them to utilize balances acquired abroad.

Now let it be supposed, the other items in the example being the same, that the imports of merchandise into A had amounted to $90,000,000. In this case gold to the extent of $15,000,000 would have been exported, giving an item of this amount on the credit side and so balancing the account. Thus we may have a favorable balance of trade and at the same time an outflow of gold. That this situation is by no means impossible in real life is seen from the fact that in 1928 the United States had a favorable balance of trade amounting to $1,038,000,000 and a net credit from gold transactions of $272,000,000.

Clearly it is the balance of payments of a country which determines the direction of the gold flow and not the balance of trade. To determine the balance of payments we must include all the debit and credit items entering into the international balance sheet except the import or export of gold itself, which is the balancing item. If

Estimated Balance of International Payments of the United States 1928

(in millions of dollars)

	Credits	Debits
Commodity trade		
Merchandise exports and imports	5,129	4,091
Silver	87	68
Ship chandlering, sales of bunker coal and oil, etc.	98	62
Unrecorded parcel-post shipments	20	20
Adjustments in above items		256
Total commodity trade (as adjusted)	5,334	4,497
Miscellaneous invisible items		
Freight payments and receipts	143	227
Tourist expenditures	168	782
Ocean-borne passenger traffic	89	..
Earnings of long-term private investments	817	252
Earnings of short-term interest and commissions	65	107
Immigrant remittances	28	217
War-debt receipts of United States Treasury	210	....
Other United States Government receipts, payments, foreign representations here	53	110
Missionary and charitable contributions, etc.	. .	52
Motion-picture royalties	70	6
Insurance transactions	80	70
Miscellaneous	46	53
Total miscellaneous items	1,769	1,876
Movement of private long-term capital		
New American investments abroad (net)	.	2,070
Reduction of previous American investments abroad	853	...
New foreign investments in the United States	1,704	...
Reduction of previous foreign investments in the United States	. .	1,223
Total private, long-term capital items	2,557	3,293
Movement of short-term capital		
Net change in international banking accounts	.	226
Pure cash items		
Gold shipments	561	169
Changes in ear-marked gold*	68	188
Total	629	357
Grand total, all items †	10,289	10,249

* Gold held in the vaults of a bank, but segregated and marked in such a way as to show it is the property of another bank.

† It is apparent that the credits and debits do not exactly balance, but it should be realized that many of the figures represent estimates only, and necessarily so, since there is no law compelling tourists to inform the government of the sums they spend abroad or compelling American shippers to certify as to the amounts they pay foreigners for ocean freight. It must also be remembered that there is no settlement date upon which a country's account for international payments must be balanced. The balance of payments is a continuing account, and even if our statistics were exact to the last dollar we should not expect an exact balance.

the sum total of the debits exceeds the credits, gold will have to be exported to bring the accounts into balance, and in this case we may say that the balance of payments is unfavorable; if the total credits exceed the total debits, gold will have to be imported to balance the accounts, and the country is said to have a favorable balance of payments.

Limits to the flow of gold. Returning to the example of countries A and B dealing exclusively with each other, let us suppose that A in a given year imports $100,000,000 in gold because of an extraordinarily large favorable balance of payments. The gold comes into the hands of bankers in the first place and either is gradually put into circulation or serves as a reserve for other forms of currency. Sooner or later its influence is likely to be felt in credit expansion, increased business activity, and a rise in the general price level. Similarly the withdrawal of $100,000,000 from circulation in B may cause, other things remaining the same, a restriction of credit, a slackening in business activity, and a fall in the general level of prices.

While gold was flowing into A the rate of exchange must have been down to the gold importing point. Now however there comes a transformation in the exchange market. The rise in the price level in A will make A a better market for the sale of goods from B but a poorer market from which to buy. Similarly B will become a cheaper market from which to purchase and a poorer market in which to sell. Imports from B into A will therefore tend to increase, while exports from A to B will tend to decline. The increase in imports means an increase in the demand for bills of exchange, while a decrease in the supply is caused by the falling off of exports. The simultaneous increase in demand and decrease in supply will have the effect of raising the rate of exchange to some point above the gold importing point in A, thus putting a stop to the further inflow of gold.

Likewise an outflow of gold from country A to country B would be only temporary, for the export of gold would reduce the price level in the gold exporting country and raise the price level of the gold importing country. In the first country the export of goods and services would be stimulated and the imports checked until

the increase in the supply and the decrease in the demand for bills of exchange would cause the rate to fall below the gold exporting point, so stopping the outflow.

We conclude that gold cannot flow in one direction, either into or out of a country, in large amounts for an indefinite period, because the flow of gold provides its own automatic check through its effect on the price level, the exports and imports, and the rate of exchange.[1] Temporary discrepancies in the balance of payments are settled from time to time in gold, but economic forces work in the direction of reducing the flow of gold to the minimum, and while in a given year a country may have considerable net imports or net exports of gold, consideration of a longer period will show the imports and exports of gold tending approximately to offset each other. The tendency over a period of time is for the debits and credits -- the amounts owed by residents of a country to foreigners and the amounts owed by foreigners to them — to equal each other. We pay foreigners for goods and services by the export of our own goods and the rendering of services to foreigners; in the last analysis the transfer of gold becomes trifling.

This result cannot be altered by artificial measures, such as protective tariffs. They have power only to change the amount of either exports or imports, which must then be paid for by means of foreign exchange in the regular way, with the same tendencies toward equilibrium. This disposes of the second fallacy in the Mercantilist doctrine.

A gold producing country. An exception to the general conclusions reached in the previous section appears in the case of a country which is normally devoting a considerable share of its energies to the production of gold. If A were suddenly to discover rich deposits of gold and were to start mining on a large scale, the first effect would be a rise in the price level. The higher prices in A as compared with B would, following the principles with which we are now familiar, lead to an increase of imports and a decrease in ex-

[1] The automatic operation of this check can of course be nullified if the country receiving the gold "sterilizes" it; *i.e.*, prevents it from entering, directly or indirectly, into circulation and consequently from affecting the price level, and if the country losing the gold neutralizes the effect of this on its own price level by a more liberal extension of bank credit.

ports, and the exchange rate in A would rise to the gold exporting point and tend to remain there. Gold would normally and regularly be exported in payment for goods imported from B. Similarly B would year after year have large imports of gold. Gold, while still the balancing item in the commerce of the two nations, should be considered in the light of a domestic product. Country A happens to have advantage in the production of gold which leads her to devote her energies to the exploitation of the gold fields rather than, let us say, to the growing of wheat, and she finds it more profitable to secure her imports by the exportation of gold than by the exportation of wheat or some other commodity in the production of which her comparative advantage is less.

Reaction of exchange rates upon imports and exports. Aside from the regulating force of the international flow of gold, the exchange rates themselves react upon imports and exports and tend to bring about an equilibrium. This reaction will become clear if, placing ourselves in the position of the merchant whose business brings him into the exchange market as a buyer or seller of bills, we consider the effect of the exchange rates upon his transactions. Let us assume that this typical merchant is exporting to England and that sterling bills are selling at a high premium in New York. With exchange in this position, the merchant may make an extra profit from his sales in England. His normal trading profit is computed in the sale price of his goods and is therefore included in the face of the sterling bills which he draws against the English buyers. In turning his bills into dollars at the prevailing high rate of exchange, a second profit is derived from the fact that each pound sterling is worth more dollars than normally. Now if this merchant is in active competition with other exporters, as is the normal case, the profitable market for American goods in England will result in lower sale prices, to the benefit of English buyers, and this situation will increase exportation, since it is axiomatic that more can be sold at a low than at a high price.

Obviously the position of the American importer of English goods is the reverse of this when sterling exchange sells at a premium in New York. When payment for goods imported from England is made by the purchase of sterling bills in New York for remittance

to the English exporter, the American importer finds his profit cut down by the premium on exchange. The premium becomes an additional cost in the business of importing English goods, and the prices of these goods will rise and their import diminish.

These illustrations should suffice to make it clear that the relation of the rates of exchange to the balance of payments is one both of effect and of cause. The rise and fall of the rates are controlled fundamentally by the balance of payments. But when an excess of imports forces the price of the foreign monetary unit (*i.e.*, the rate of exchange) to a premium, the high rate tends to restrict imports and encourage exports and thus works toward a restoration of equilibrium in the foreign trade. Conversely when a large volume of exports makes the rates on foreign bills fall to a discount, imports are stimulated and exports restricted, which again tends to bring the trade into balance. In normal times the rates, confined as they are within the bounds of the gold points, can rise or fall so little that their influence upon the balance of payments may be imperceptible. It must be borne in mind however that this narrow range exists only because gold is permitted to flow in sufficient volume to offset the balance of payments; when gold is prevented from supplying this corrective, the rates of exchange will exert an influence tending to restore equilibrium.

Is gold a peculiarly desirable import? Many persons would immediately answer this question in the affirmative. They experience a feeling of satisfaction when the statistics show that gold is coming in, they "view with alarm" an outflow of gold, and they would favor governmental action to encourage the import of gold. This is the third of the Mercantilist fallacies to which we referred in the opening paragraph of this chapter.

In the first place, it should be evident that gold obtained through international trade adds nothing to the total wealth of the nation, since it is obtained only by the sacrifice of other goods presumably of equal value. Even if the gold were obtained by an international loan, a debt is created equal to the value of the gold obtained. Those who believe gold imports are peculiarly favorable must base their belief upon the notion that gold is in some way superior to other forms of wealth or property. Is there any basis for this

notion? Gold has certain important industrial uses, as in dentistry and the making of watches and jewelry; its principal other use is to serve as a medium of exchange. So far as its industrial uses are concerned, it will scarcely be maintained that gold is more essential to a people than a host of other commodities — wheat, meat, coal, iron, steel, lumber, and so on indefinitely. If a country lets these other things go in order to get gold, it presumably gains in utility through the import of gold. But it just as truly gains when it exports gold in order to obtain other things which it wants more. Gold differs not at all from other commodities in this respect, and the student who has any knowledge of economic principles need pursue this part of the inquiry no further. Certainly nothing could be more absurd than, by governmental action, to compel a people to give up things they want more in exchange for gold they want less.

The volume of monetary gold. As regards the use of gold as money, the important fact is that, within limits, the absolute quantity of money is pretty much a matter of indifference. Through the adjustment of prices, a given volume of exchanges may be consummated with either a small or a large quantity of money; what occurs is that the price level is low in the one case, high in the other case. To acquire, by artificial means, a greater quantity of gold would raise the local price level, disturb the relations between debtors and creditors, disorganize the international balance of payments and the exchange market, and finally, unless prevented by other measures equally artificial, lead to the expulsion of the gold itself, all to no purpose. Also disastrous is the effect of this on the economic life of the countries from which the gold is artificially withdrawn, where in the absence of countervailing measures a declining price level, decreased business activity, and all the other maladjustments of a deflationary period are the natural consequences.

If the gold standard is in general international use, if the available supplies of gold are equitably distributed among the nations in rough proportion to their currency needs, and if internal prices in the various countries are in a condition of substantial equilibrium with each other, the movements of gold, whether in or out, in response to changes in the balance of payment should occasion

neither rejoicing nor alarm. But in recent years none of these conditions has obtained. The United States has acquired a disproportionate share of the world's gold, and the internal price levels of the countries of the world, thrown out of adjustment by war and post-war conditions, have not yet been brought into stable equilibrium with each other. One by one nations have been forced or have found it expedient to abandon or modify the gold standard. When a country is trying to maintain the gold standard in the face of a rapid depletion of its stocks of gold, the export of gold in any considerable amounts is a matter of utmost concern to the treasury and bank officials of the nation and justifies, so people think, any restrictive measures which may so operate on the international balance of payments as to prevent further export.

Likewise a country which is contemplating a return of the gold standard in one form or another finds the acquisition of gold reserves essential and will undoubtedly consider gold so valuable an import as to warrant substantial sacrifices in goods and services. The resumption of specie payments after the paper money inflation of the Civil War was rendered easier by a turn in the balance of payments in favor of the United States. An unusually large wheat crop coupled with disaster to the foreign crops came in time to check the outward flow of gold and start it flowing the other way. For twenty years thereafter the maintenance of the gold standard, in the face of inflation by greenbacks, silver dollars, and Treasury Notes of 1890, was a more or less heavy burden upon the Treasury and gave to the international movement of gold an interest which, except for the unsound monetary system, it need not have had.

In general however there is nothing, apart from the needs of a monetary system at critical periods, to set gold apart as a peculiarly desirable import and nothing to be gained by forcing its import in excess of what would enter under the normal operation of international trade. Under conditions such as existed prior to the World War the flow of gold is an automatic thing, controlled by economic forces which result in a state of equilibrium which is for each nation the most favorable situation possible in view of all its other circumstances.

Is an excess of imports a sign of national decline? We have seen that a favorable balance of trade does not necessarily mean an inflow of gold, that an unfavorable balance of trade does not necessarily imply an outflow, and further that even if it were possible to bring about the importation of gold by artificial means, the final result could only be a rise in the price level and a decline in the purchasing power of the gold unit. However in view of the present-day persistence of Mercantilist theories and the rather common dread of an unfavorable balance of trade, we may be justified in giving specific consideration to the popular notion that a country which has an unfavorable balance of trade is losing its vigor and is on the road to economic ruin.

England, Germany, and most of the highly industrialized countries have had unfavorable trade balances for many years. Year after year these countries have imported a greater volume of commodities than they have exported. The United States, on the other hand, has long had a favorable balance of trade. Each year a portion of her exports has gone to pay for the use of foreign capital, for shipping, and banking, and other services rendered to Americans. A decrease in America's favorable balance of trade indicates that her dependence on those services is becoming less, and an unfavorable balance would mean that foreign countries were becoming dependent upon her, were paying for these services with their manufactured goods or raw materials. An unfavorable balance of trade is far from being a stigma of decadence; it simply indicates that the economic relations are different from those of a country with a favorable balance of trade. In particular it is apt to be the mark of an old established nation with a large accumulation of wealth.

EXERCISES

1. Construct an international balance sheet from the following items, and determine the balance of trade and the balance of payments.

Exports of merchandise	$2,400,000,000
Shipping services rendered by foreigners	80,000,000
New American investments abroad	100,000,000
Interest received by American investors	65,000,000
Imports of merchandise	2,100,000,000
American tourist expenditures abroad	320,000,000

New foreign investments here	600,000,000
War debt receipts	5,000,000
Immigrant remittances	35,000,000
Foreign tourist expenditures here	80,000,000
Interest paid to foreign investors	75,000,000
Gold	....?

2. Explain how each of the following tends to affect the price of foreign exchange in this country:

(*a*) Decreased export of American cotton
(*b*) Development of an American merchant marine
(*c*) Stoppage of war debt payments to this country
(*d*) Decline of immigrant remittances
(*e*) Increased imports of Canadian lumber
(*f*) Stoppage of interest payments due American investors
(*g*) Redemption of American securities held abroad
(*h*) Loss of confidence in the stability of a foreign currency, causing foreigners to increase their investments here
(*i*) Increased sales of American automobiles abroad
(*j*) Decline of American tourist expenditures abroad

XXXVI

CONTEMPORARY PROBLEMS OF FOREIGN EXCHANGE

Pre-war equilibrium. In the two chapters immediately preceding, we have followed the workings of foreign exchange as it prevailed in the period before the World War, when the important nations were on the gold standard. It was a period not only of stable exchange relations but of substantial equilibrium as well. Changes of course were taking place, and there were periods of excessive business activity and periods of recession, but there was a mutual adjustment of fundamental factors in national and international life which we do not have today. Price levels were in equilibrium. International trade followed lines determined by a fairly well-settled international division of labor and was not so constricted as today by universal high tariffs and other nationalistic restrictions nor so disorganized by frequent changes in tariff and trade regulations. International investment proceeded along conservative lines under the leadership of the experienced investment houses of London.

This international equilibrium was brought about largely by the normal operations of the gold standard, which provided, as we have seen, automatic checks to factors which might make for disequilibrium. *Abnormal factors operating through the price level or the* balance of payments could be taken care of by exchange operations without causing disaster either to the exchanges or to the internal economy of a country. The operation of these automatic controls tended to keep the rates of exchange normally within the gold points, but it should be noted also that central banks often intervened in the market to bring the rate of exchange down from the gold exporting point when it was believed that this rate was the result of temporary dislocations in demand and supply. The standard method was for the central bank to raise its rate of rediscount.

A rise in the rate of rediscount at the Bank of England, for example, would make it more profitable to leave funds in London than to transfer them to New York, or it might even induce funds to come there, thus altering the demand and supply of bills of exchange on that centre in such a way as to check the outflow of gold or to reduce its volume.

Post-war exchange conditions. It is unnecessary to emphasize the internal and external dislocations brought about by the World War, but it is important — even though we can do no more than mention a few of the major ones — to show why these dislocations and others, the product of the post-war period, have given the world its heritage of disordered exchanges. In the first place, it is clear that the practical problem of financing war needs made both the abandonment of the gold standard and some degree of inflation almost imperative. In the first stages of the war, gold coin was withdrawn from circulation in the warring countries, and an embargo was placed on the free export of gold, thus removing one of the limits to exchange fluctuations, the gold export point. This meant that in New York sterling exchange could not rise above the gold export point, for Americans were still able to export gold, but it could fall below the gold import point (the British gold export point), for gold could not be freely exported from Britain. As a matter of fact Great Britain found it desirable at first to part with a large portion of her gold reserves in order to prevent the exchange rates in New York from falling too rapidly and too far, and eventually it seemed expedient to her and to France to "peg" the rates at a fixed point. Britain effected this by arranging with J. P. Morgan and Company to undertake to buy all sterling exchange offered at a fixed price of $\$4.76\frac{7}{16}$ per pound. Private loans raised in America, *funds secured from the export of American securities held abroad,* and finally advances from the United States Treasury supplied the funds for these pegging operations.

Inflation came about as a result of unbalanced budgets and was resorted to in degrees which varied with the financial strength of the governments concerned and their ability to raise revenue either by taxation or by bond issues. It meant of course the printing of inconvertible money and the expansion of deposit currency through

government borrowing from central banks for the purpose of meeting current governmental expenditures. At the time it was in some cases justified as an emergency measure, and subsequently the inevitable rise in price levels has frequently been justified as a method of making the burden of the national debt tolerable. In Germany inflation proceeded to such an extent and prices rose to such a fantastic level that the national internal debt was virtually eliminated.

The United States avoided the necessity of balancing its budget by the emission of paper money, but it did not escape inflation, which was occasioned by the importation of millions of dollars in gold and by increases in federal reserve notes (a comparatively new form of money at this time) and bank credit made possible by the influx of gold.

Two facts of special significance emerge from this monetary disorder. The first is that exchange rates became erratic and subject to violent fluctuations as soon as government support was withdrawn, leading to a vast increase in the risk entailed in foreign transactions and to an unaccustomed volume of speculative operations in the exchanges. The second is that inflation proceeded in different degrees in the inflating countries, so that the natural relation of price levels to each other was distorted.

Preoccupation with war and with the production of war materials led to the growth of huge war industries in the European countries and to the decay of other industries, notably of some which had been most successful in the export trade. Lack of shipping space, as well as a lack of goods for export, transferred much of the export trade of European countries to the United States. The creation of new national boundaries by the peace settlement separated regions which had been economically unified, and this dislocation was intensified by an increase of tariff barriers which a new and fiery nationalism demanded. International trade was thrust from its accustomed channels and impeded in ways which brought disastrous results on the balance of payments of various countries.

Finally the heritage of war debts between the Allied and Associated Powers and the enormous reparations payments demanded by the Allies added another strain to international finance, a strain

intensified and indeed made intolerable by the unwillingness of many countries to receive payments in goods and services.

Irredeemable money and the par of exchange. Before discussing the problems of the exchanges in the post-war period we must understand the relation existing between the money of a country on an inconvertible paper standard and that of a country on a gold standard, and also the relation between the respective units of two countries both of which are on an inconvertible paper standard. When the British government in 1919 "unpegged" sterling exchange in New York; that is, when it ceased to give it an artificial price by a continuing offer to buy sterling at a fixed rate, it was said that sterling was free to find "its natural level." Today it is said that Great Britain cannot return to the gold standard until the pound has found its natural level in terms of the dollar. What is a "natural" rate and why is another rate an "unnatural" one?

The purchasing power par of exchange. The answer to this question is found in what is called the purchasing power par of exchange, which may be defined as a rate of exchange between two currencies which expresses an equivalence in the purchasing powers of the respective units in the home countries. Thus if $6.00 were to be assumed to be the purchasing power par of exchange between the dollar and the pound when both the United States and Great Britain were on a depreciated paper standard, it would mean that the pound had six times as great a command over commodities in general in Great Britain as one dollar had in the United States.

The purchasing power par of exchange is of course not a fixed par, but fluctuates with changes in the purchasing powers of the respective units, as measured by changes in the index numbers of price levels. If the price level in the United States were to rise while the British price level remained unchanged, the purchasing power par would move upward from our assumed figure of $6.00, while if the British price level were to rise and the American price level to remain stationary, the purchasing power par would decline from $6.00.

Perhaps the most important characteristic of this par of exchange is that it is neutral in its influence on the currents of commerce. It is a rate which would neither stimulate nor check imports or exports but would permit them to pursue their normal course. It would

neither discourage nor encourage capital imports or exports; it would neither bring about an influx of tourists nor would it serve as a bar to their coming.

Purchasing power par and market rate of exchange. The purchasing power par of exchange is always a useful concept in any period of economic equilibrium and when we are considering two countries whose trade with each other is sufficiently large to influence the general level of prices in each country. Market rates of exchange will actually tend to revolve around this neutral or equilibrium par of exchange. A rate higher than our assumed par of $6.00 would mean that sterling exchange was at a premium in New York. Imports into the United States would consequently be checked and exports would be encouraged. An increase in the supply of sterling exchange and a decrease in the demand in New York would cause the rate of exchange to fall toward the purchasing power par of exchange. Similarly a discount on sterling in New York would affect the currents of commerce in such a way as to bring the rate of exchange up toward the purchasing power par.

In a period of rapid economic change however the adjustment of prices to disturbing factors is very slow, both nationally and internationally, and in such a time the market rates of exchange may bear little direct relationship to the purchasing power par. Let us suppose for example that the United States should inflate its currency. The immediate effect would be a rise in the rate of foreign exchange; it would take more dollars to buy a pound sterling for example. Eventually there would also result a corresponding rise in the prices of American goods. If this effect also came instantaneously, the monetary inflation would have no influence on export trade. A pound sterling would cost more in American dollars, but American goods would be worth more in dollars, so that in terms of domestic goods foreign exchange would be no more costly than before, and there would be no tendency to check imports. From the foreigner's side, American dollars could be obtained more cheaply in terms of, let us say, British pounds, but American goods would be correspondingly higher priced in dollars, so that there would be no added inducement to buy American goods; American exports would therefore not be encouraged.

Actually it takes considerable time for the general price level to adjust to monetary inflation. In the meantime therefore the high exchange rate, caused by monetary inflation, exerts its normal influence to check imports and stimulate exports. The effect of this on the rate of exchange, other things being equal, would be to check the speed with which it was rising. None the less, so long as inflation proceeded, we would expect the rate of exchange to be continuously at a premium over the purchasing power par of exchange. When inflation ceased and the domestic price level had adjusted itself to the new quantity of currency in circulation, we should expect the rate of exchange again to be determined around the purchasing power par.

The actual situation in a period of inflation is of course by no means so simple as this, for the lack of confidence engendered by inflation may affect many items in the balance of payments and through them the rate of exchange. There would for example very probably be an exodus of capital, a "flight from the dollar," to countries in whose monetary stability there was greater confidence. This would bring so much pressure on the exchanges as to force the rate up even more rapidly than would otherwise be the case, and in turn the repatriation of this capital in the days of returning equilibrium would have a depressing influence on the rate.

The purchasing power par of exchange is today a matter of practical importance. To countries contemplating a return to the gold standard it is essential to discover if possible this neutral rate of exchange; to officials charged with the duty of eliminating excessive fluctuations in the market rates of exchange it is desirable to know what figure may be taken as representing a natural rate of exchange between two countries; and in periods of rapid change it is useful to know the direction in which this par of exchange is moving, even though the causal relationship between the purchasing power par and the exchange rate may be obscured by the operation of other factors.

Unhappily the practical determination of the purchasing power par of exchange is beset with apparently insuperable difficulties; at least thus far no computing method has been discovered which is accepted as theoretically sound. The experimental method is the

only one available, and it is the one which must be used by officials in charge of exchange stabilization funds. On the basis of indexes of price levels and of all other pertinent economic data, some rate must be chosen as a tentative par of exchange. The officials must then by appropriate operations in the exchange market strive to keep the market rates of exchange from diverging too widely from this assumed par of exchange. If it is found that as a result of these operations the market rates are continuously discouraging exports, it is clear that they are below the purchasing power of exchange, and the officials must revise their tentative par of exchange. By a series of shifts of this character they may hit upon the purchasing power par, but even so it should be noted that their difficulties are not at an end, for the par is or may be a moving par.

The stabilization of exchange rates: General advantages. In a period of economic disequilibrium unstable exchange rates are a usual characteristic of currencies divorced from gold, both because the par of exchange may fluctuate with changes in the price levels of the countries concerned and because the fluctuations of the rates of exchange around this par are not subject to any such limitations as are imposed by the gold points.[1] Even in periods of relative stability, when the par of exchange moves very slowly or not at all, temporary influences on the exchanges may bring about a large premium or discount in the rate.

Unstable rates of exchange are a source of much disturbance in world finance. Not only do they add to the risk of international trade and thereby tend to diminish its volume, but they induce speculative dealings in exchange which may themselves cause disturbance. Hence most countries are interested in reducing the range of fluctuations in exchange rates or in stabilizing the foreign prices of their currencies at some given level, unless indeed they are themselves inducing a series of fluctuations in order to bring the rates to a higher or lower level. Stabilization plans may contem-

[1] It should be observed that when one country is on the gold standard and another on depreciated paper there will be gold points so long as gold may freely be shipped. Gold will be quoted in paper dollars just like any other commodity. Consequently the cost in paper dollars of purchasing the fine gold content of a foreign standard coin plus the cost of remitting this to the foreign country will determine the maximum rate which would be paid for a bill of exchange on that country. This is the gold export point. In like manner the gold import may be determined.

plate the return to the automatic controls provided by the gold standard, or they may involve various types of exchange control which have developed in recent years.

The reader is already aware of the many practical and psychological advantages to be obtained from a return to the automatic controls of the gold standard, augmented as might be desired by other methods of control. At the close of the World War these advantages appeared so compelling that nearly every country looked forward to the time when it could feel it was in position to resume the redemption of its paper money in gold and permit the free export and import of gold.

Many countries succeeded in this during the ten years following the war. We can do no more than touch briefly on the history of stabilization in Great Britain and France, whose experiences in returning to gold are instructive in view of the fact that this problem confronts so many countries today.

Stabilization of the pound. Partly for reasons of national pride and partly because of pressure from the class of investors in government securities, Great Britain returned in 1925 to the gold standard at the old gold content of the pound. Thereafter the gold or mint par of exchange of the pound to the dollar was much higher than the natural level of the two currencies, as determined by the purchasing power parity. And the consequent overvaluation of the pound in terms of the dollar introduced serious elements of instability into British economy. At this time also British public finance was in difficulties. Heavy increases in government expenditures, with inadequate taxation, were causing deficits which contributed to loss of confidence in the financial stability of the British government. These influences, together with other disturbing factors, undermined and finally overthrew the gold standard.

To understand the influences of the monetary factor, we have only to consider the effect of the new level of rates on exports and imports. Assume that a fair degree of stability had been achieved in the internal and external relations of Great Britain prior to revaluation, and then ask what effect the new exchange rates would have on exports and imports Clearly they would raise the prices of all British export goods in foreign markets and decrease their sale;

quite as clearly the price in Britain of foreign goods would decline and there would be a tendency for imports to increase. This would in turn force the rate of exchange, in the absence of counteracting factors, toward the gold export point, and if gold actually were to flow out, it would tend to restrict the extension of credit and to raise the price of credit. Depression in the export industries and in industries facing increased competition from imports could not fail to have a depressing influence on the whole industrial fabric. In other words, the result of stabilization at too high a point is at best a gradual retarding of business activity, made more difficult by rigid wage and interest rates; at the worst it means depression and possible abandonment of the gold standard.

The impact of these forces was softened for the moment and the seriousness of the situation in Great Britain obscured by the operation of a number of factors, one of which requires special mention, for it has been a plague to all nations. This was the growth in the volume of liquid funds seeking profitable short-time investment in one market or another as the rate of interest might swing up or down, or looking for an opportunity to speculate in exchange rates, or more frequently seeking safety without much regard to profit in the market which seemed for the moment to be most stable. The return to the old gold parity had induced in financial circles an unwarranted belief in the stability of London, and a large volume of liquid capital had been transmitted there for shelter and temporary investment. This obscured the real pressure on the exchanges, because the mechanics of transferring these funds to London served to counteract the tendency of changes in the currents of commerce to send the rate to the gold export point. They were however a constant source of weakness, for at the first hint of financial difficulties they would be (as later they actually were) transferred to another haven.

In summary we may say that the attempt to restore the gold standard at a mint par of exchange substantially above its purchasing power parity subjected the exchanges to a continual strain and in addition exerted a depressing influence, not only on the export industries, but on all other industries as well. An unstable equilibrium in the balance of payments was achieved, largely by an

influx of short-time investments, ready to move out at the first sign of danger and not counterbalanced by equivalent short-term investments by London in markets whence they could be repatriated. London did, it is true, have short-term investments in large quantities, but when she needed them they were "frozen" in Germany, Austria, and elsewhere. When therefore the European crisis was precipitated in the spring of 1931 by the failure of the Credit Anstalt of Vienna, and panic spread from Austria to Germany and then to Great Britain, gold began to be withdrawn in large quantities as panic-stricken investors transferred their funds to more stable markets. Despite active intervention in the exchange market facilitated by large loans from the United States and France, it was impossible to hold on longer to the gold standard. In September, 1931, Great Britain was again on an irredeemable paper standard.

Stabilization of the franc. After the World War France experienced a stormy period of inflation, which was marked by depreciation of the internal and external values of the franc, wild fluctuations in the exchange rates caused by the depreciation of the currency and excessive speculation in exchange, and a flight of capital from the franc. Ultimately it was decided to adopt the gold bullion standard and to stabilize the franc in terms of gold.

The experience of France illustrates the consequence of adopting a mint par of exchange lower than the purchasing power par. In its immediate effects France found this beneficial for, aided to be sure by tariff increases and quota restrictions on imports, she achieved a short-lived favorable balance of trade. A large favorable balance of payments, resulting in part from the favorable trade balance, in part from reparations payments, partly from tourist expenditures, and partly from the repatriation of capital which had fled the country during inflation, caused a huge inflow of gold. French industry was well occupied and the French worker and peasant reasonably contented.

The conclusion that France was clever and England foolish, as drawn by some, does not necessarily follow. While the unfortunate consequences to England are clear enough, it should likewise be evident that the gold drawn to France was taken from other countries, that the exports artificially stimulated by the rate of exchange

may have competed unduly with the production of other countries, that the restriction on imports caused by the exchange rates or by tariffs and other measures may have had a depressing influence on business activity elsewhere, and that the repatriation of French capital from countries where it had long stayed may have helped to undermine the financial structures of these countries. Were France to have no international connections, and were her prosperity to be in no wise dependent upon a prosperous world, the chain of events set in motion by her revaluation of the franc at "too low" a figure might be looked upon as desirable from a national point of view, however disastrous it might be internationally. Events of the past few years however point to an inevitable relation between national and international well-being.

Difficulties in returning to the gold standard. There would seem to be two indispensable prerequisites to a successful return to the gold standard. The first is an ample gold reserve behind the national currency, and the second is the determination of the new gold par of exchange at or very near the purchasing power par of exchange. Unhappily few nations consider their supplies of gold adequate, and in the absence of international coöperation in the redistribution of gold this situation is likely to continue. Furthermore, as we have seen, the calculation of the purchasing power par of exchange offers great practical difficulties, which in a period of rapid change become even greater, for the par itself may change from day to day. These facts lend support to the conviction of many that it is unsafe for any one nation alone to return to the gold standard and that it would be dangerous even for nations acting in concert to attempt this until a sufficient degree of stability has been reached nationally and internationally to enable the authorities to select gold pars of exchange which would not themselves become causes of disequilibrium.

Whether because of fear of the practical difficulties involved or because of belief in the superior flexibility of a monetary system which is independent of gold, few nations which have left the gold standard in recent years have returned to it in any form. Most nations have felt however the need of some form of official supervision over exchange rates or active control of exchange operations.

The problem of exchange control: Two principal means. Among the present diverse forms of exchange control we may distinguish two main types. The first involves active dealings in foreign exchange, constantly or intermittently, by national treasuries, the central banks, or special boards constituted for that purpose. The intention usually is to eliminate undue fluctuations in rates caused by temporary variations in the demand and supply of bills of exchange or by excessive speculation. Occasionally exchange operations may be intended to raise or lower the par of exchange and may form a part of a more comprehensive plan of monetary management.[1]

The second type of control is intended to prevent the transactions of international trade or finance from exerting a dangerous influence on the balance of payments and hence on the monetary system or the financial structure of the country. This may be achieved by a variety of restrictions on the freedom of the individual in foreign trade and finance or the monopolizing of all dealings in foreign exchange. Frequently also fluctuations in rates of exchange are eliminated, at least legally, by the establishment of an official rate of exchange at which all dealings in foreign exchange must take place.

Elimination of temporary fluctuations in exchange rates. The student will recall that the exchange rates of a country on an inconvertible paper standard have no limit to fluctuations such as the gold points offer and that the possibility of such variations adds considerable to the risk of financing exports and imports and hence to the cost. While it is true that in most markets, by making contracts to buy or sell exchange in the future, the trader can transfer the risk from himself, he does so only by adding the cost of this transfer of risk to the price of the good in which he is dealing. There is a further fear that temporary influences on rates may, if continued long enough, have a permanent effect on the currents of commerce and therefore affect the true par of exchange.

To guard against what seems to be an unwholesome range of fluctuations, a country may actively intervene in the exchange market. Intervention for this purpose is best illustrated by the

[1] See Chapter XXVIII.

Exchange Equalisation Account of Great Britain. This was established in 1932, with resources of £175,000,000, later increased to £575,000,000, and with power to deal in foreign bills of exchange, to buy and sell gold, and to hold balances in foreign banks. If the London rate in New York is declining below the point thought desirable, the officials can sell dollar exchange which they have acquired. A rise of the sterling rate in New York above the desired point can be counteracted by the purchase of dollars in London. By such operations the Account can throw its influence on the demand or the supply side of the market as it desires and can usually exercise a compelling influence on the movement of the rates.

Attention should perhaps be called to two difficulties in the operation of a stabilization fund. The first concerns the choice of the rate of exchange for purposes of stabilization; the second relates to the ability of exchange officials to distinguish between those movements of rates which are caused by temporary variations in the supply and demand of bills of exchange and those which reflect a change in fundamental conditions underlying the balance of payments. Errors of judgment in these two directions may serve to augment rather than to smooth out difficulties. Nevertheless the belief that it is wiser to run the risk of occasional mistakes in judgment than to leave the exchanges entirely uncurbed appears to have been confirmed by the success of stabilization funds in securing relative stability of rates.

Intervention to curb speculation. We have already spoken of the large volume of liquid capital which is available for temporary investment in one market or another as exchange profits or relative safety dictate, and we have pointed out the dangers which are involved in this. At one time during the American experimentation with devaluation it was said that over a billion dollars of capital had sought refuge in London, much to the embarrassment of London. An equalization fund can do something to mitigate this danger to the exchanges at the time the inflow of capital starts by investing some of its resources in foreign exchange, which can later be thrown into the market when a change in foreign exchange conditions makes it profitable to send this capital to another centre. It can, in other words, sterilize this influx of capital so far as its influence

on exchange rates is concerned, provided the resources of the fund are adequate. If the inflow is too great for the resources of the fund, then resort to restrictive measures may be the only expedient at hand.

Speculation in the exchanges takes of course other forms than transfer of timid capital from one centre to another. In each market there are individuals dealing both in "spot" and "forward" exchange who are expecting to profit by their correct anticipation of the movements of these rates and who by their speculative dealings aid in keeping a proper spread between the present and future prices. There are also those who make a business of dealing simultaneously in the exchanges of several different countries and by so doing keep them in adjustment one with the other. Thus if the rate in New York on London, the rate in New York on Paris, and the Paris rate on London are not in adjustment these so-called "arbitrage" operations will at once be set up, tending to bring them together again. In any exchange market of consequence these speculative activities go on as a matter of course and are generally conceded to exert a beneficial influence.

Between countries solidly on a gold standard, speculative activity causes little anxiety, but in exchange markets where wide swings in the rates are possible and where there may be a moving par of exchange, speculation may proceed to such lengths as to become a basic factor in determining the currents of commerce and hence the par of exchange. It is to curb such activities that the British, French, Italian, and other exchange authorities have on occasion intervened in the market, with disastrous results either to the speculators or to the intervenor. Unhappily the just suffer with the unjust; equally true is it that fundamental trends may be interpreted as speculative activities. The use of an equalization fund for this purpose certainly presumes on the part of the fund officials an extraordinary degree of acumen, decisiveness, and technical skill in manipulating the exchanges.

Restrictions on dealings in exchange. Restrictive measures may take many forms and may affect the exchanges directly or indirectly. Among those which have been used in one country or another in the last few years we find the following: a moratorium on

the payment of interest outside the country; an embargo against the export of capital; the requirement that all foreign exchange acquired by the export of goods or otherwise be turned over to the central bank or other designated exchange authority; a requirement that exchange must be arranged for with the central bank before an import license will be granted; a quota system for imports, which may serve also as a protective measure to domestic industry; the stipulation that exchange will not be available for the payment for imports from a specific country unless it has been previously acquired by the export of goods to that country.

While exchange restrictions were not unknown before the onset of the world depression (after 1929), their recent extension has been at the same time the result of conditions of depression and a cause of the continuance of the depression. A surprisingly large number of countries have today rather comprehensive systems of control — among others, Italy, France, Switzerland, Germany, Hungary, and the other countries of the Danubian basin, Turkey, and many countries in South America. Devised, as we shall see, as protective measures in troubled times, they fit rather perfectly either into plans of managed currency or into more comprehensive plans of national control of economic life.

To understand why so many countries have resorted to restrictive controls it is necessary to recall that in the post-war period a precarious balance in payments was achieved in many of them only by large loans from the United States, that their gold reserves were dangerously small and their monetary systems easily thrown into confusion, and that many of them were overburdened with interest payments, short-time investments, and debts for past importation of goods. The crisis in the United States put a complete stop to the floating of international securities in the American market, thus placing a strain on the balance of payments of all countries which used exchange acquired by these loans to meet current indebtedness. At the same time all countries found the world demand for their products declining and the world prices of their products, particularly of agricultural, mineral, and other staple goods, decreasing rapidly.

The danger to the exchanges and to internal monetary and bank-

ing structures seemed in many countries critical enough to warrant drastic measures. The most obvious move was to stop the transfer of the payment of debts owed abroad. The Hoover Moratorium of June, 1931, on the payment of interest and principal on inter-government debts and reparations set a precedent which was willingly followed by many countries.

Moratoria, expanded to postpone the payment of all indebtedness whether arising from bonded indebtedness, short-term investments, or imports, served to prevent the obligations of a more optimistic past from exerting a disastrous influence on the balance of payments, but in many countries equal concern was felt lest current international transactions have a similar influence on the balance of payments, and some were anxious to acquire a favorable balance from current transactions which would enable them gradually to liquidate their unpaid foreign debts.

To attempt to secure a balanced account or a favorable balance through an increase in exports seemed futile; the only other method was a reduction in imports. In the long run this is not only equally futile, since one country's inports are the exports of some other country, but it is positively harmful because of the influence on the volume of world trade of a series of import restrictions. Yet it was an expedient which might secure results for a few weeks or a few months before other countries adopted similar measures.

Effective command of the situation demanded a concentration of exchange control in the hands of the government, for thus only could undesired fluctuations in the rates be prevented and only in this way could all exchange acquired by exports be made available both for the payment of imports and for the liquidation of such portions of past indebtedness as seemed expedient. This type of restrictive control involves usually these three things: (1) the surrender of all foreign exchange by exporters and others to some representative of the government, (2) a system of import licenses, issued according to the availability of exchange and sometimes according to the importance of the import to national life, and (3) an official rate of exchange set by the exchange authorities. A considerable though decreasing degree of difficulty has been experienced by most countries which have attempted to monopolize

exchange transactions in this way, for coincidental with the expansion of official control came the growth of what were known as the "black bourses"—subterranean markets for "bootleg" exchange.

The balance of trade. In all countries which exercise strict control of the type just described, great importance is attached to the balance between exports and imports of goods, because of the fear of the influence which an unfavorable balance of trade might exert on the exchanges. In some countries indeed the novel doctrine has been evolved that a balance in exports and imports should be secured with each separate country. It has a specious appearance of fairness, and a surprisingly large number have been seduced by it. For example the trade diplomats of Great Britain have convinced Argentina that she should purchase more British goods and fewer American goods since the Argentine balance with Britain is favorable while her balance with America is unfavorable. Germany informs America that unless Americans buy more from her she will reduce her American imports, and recently she has done so, bringing her American exports and imports into balance. While this type of balance may offer some practical advantages to countries which have set up exchange clearing arrangements, it rests otherwise on a complete misunderstanding of the principles of international trade, according to which it is the total balance of payments, not the separate trade balances with particular nations, that affects exchange rates and the monetary situation. The final result of a forced equilibrium of exports and imports with particular nations can be only a decrease in the total volume of world trade, a decline *in interregional division of labor, and a shrinking in world income. To justify this policy as a weapon of commercial diplomacy* is to mistake the bases of world trade; exports and imports exist and increase because of the mutual advantages secured from the trade *and not because of bludgeoning methods used by one country* towards another.

Exchange clearing systems. Exchange clearing systems represent in origin a device whereby one country may secure payment for exports to another country or for interest which is due but withheld. As a weapon of this sort it is available only to a country which

has an unfavorable balance of trade with another country. In actual practice the clearing system may have been established by mutual agreement, but behind the appearance of an amicable settlement there is the possibility of forcible collection.

To illustrate this let us take the trade of Hungary and Switzerland. Switzerland was importing more from Hungary than she exported to her and was paying for her imports, whereas the Hungarian government decreed that payment for Hungarian imports must be made in national currency into a special account with the Hungarian National Bank. Assuming that the Swiss exporters were not content to have "blocked" accounts in the Hungarian National Bank, the reasonable solution was either for Switzerland to cease exporting or to impound in Switzerland the money Swiss importers owed Hungarian exporters, make a deduction sufficient to cover payment to the Swiss exporters, and remit the balance to Hungary. As a matter of fact Hungary was induced to see the reasonableness of this, and in November, 1931, an agreement was signed whereby payment for Hungarian exports to Switzerland was to be made in francs by the Swiss importer into a special account in the Swiss National Bank, and payment for Swiss exports to Hungary was to be made in pengoes by the Hungarian importer into the clearing account in Hungary. Thus the payments of Swiss importers made funds available in Switzerland for payments to Swiss exporters, and the same was true in Hungary. The arrangement included practical details for balancing the two clearing accounts, which need not concern us here.

Undoubtedly the system has practical merits in trading with a country of weak financial standing. It gives the countries effective control of their trade balances, it insures payment to exporters of both countries, *it removes the exporters and importers entirely from the exchange markets, leaving the matter of transfer to governmental officials*, and it is probably true that it *removes some of the* incentive to erect additional tariff barriers, exchange restrictions, and the like. Because of these practical advantages many of the countries which have adopted stringent restrictive measures of control over the foreign exchanges have turned to exchange clearing systems. As a part of a permanent trade policy it is of more than

doubtful wisdom, unless the goal be the control of the character of international trade and not an expansion in its volume.

Exchange control in the United States. In concluding this chapter on recent developments in the foreign exchanges, a word must be said about dollar exchange. The changes in the American monetary and banking system have been described in previous chapters in sufficient detail to give the reader some idea of the impact of these factors on foreign exchange. The extent of the changes however and the degree to which control over the exchanges has been lodged with the executive branch of the national government are perhaps not fully appreciated.

It will be recalled that after a period of depreciating dollar rates in foreign markets, the United States, in January, 1934, set up a system whereby gold might be obtained for export. with a devalued gold dollar defined as containing $15\frac{5}{21}$ grains of gold nine-tenths fine. This represents a decrease in weight of about 41 per cent from the old gold dollar. It should be noted that this is not necessarily the final weight, for Congress has given the President continuing authority up to June 30, 1939, to alter the weight by proclamation and to fix it at any point between the limits of 12.9 and 15.48 grains. We have therefore the possibility of a changing par of exchange with other countries.

A more important test of the gold standard than the weight of gold ascribed to an imaginary coin is the presence or lack of a free gold market, and a free gold market the United States does not now have. Her paper money is not redeemable in gold, either coin or bullion, and while gold bullion and foreign gold coin may be exported and imported for private domestic and foreign account, this is only under license and regulation of the Secretary of the Treasury. Gold does not therefore move in or out automatically as it does in a country on a gold standard. There are, to be sure, points in the exchange rates with gold standard countries, but they are not certain in their control of imports and exports of gold, for they exist at the pleasure of the Secretary of the Treasury.

A further development has been the establishment in the Treasury of a stabilization fund under the direction of the Secretary of the Treasury. This fund was established by the gold reserve act of

1934 and was provided with $2,000,000,000 of the $2,811,375,757 profit secured from the revaluation of gold. In the original act the life of the stabilization fund was limited to a maximum of three years, but by an amendment passed in January of 1937 this has been extended to June 30, 1939. In its objectives and operations it is quite similar to the British Account. The Secretary of the Treasury is empowered to deal in gold and foreign exchange and to conduct other operations which will stabilize the value of the dollar exchange in foreign currencies and maintain the market for government securities. It is obvious that the operations of this fund are to be directed against undesired fluctuations in the rates of countries on an irredeemable paper basis. This might conceivably result in a clash between the operations of national funds. Manipulations of the British to depreciate the sterling rates because of the encouragement to their export trade might be offset by the operations of the United States Secretary of the Treasury if he felt that the dollar rate was too high in terms of sterling and other currencies. The result might be a stalemate or a verdict awarded to the largest fund — neither of which would tend to improve international amity.

Possible danger from competing stabilization funds has been banished, for the time being at least, by an agreement reached in the fall of 1936 by the United States, Great Britain, and France, and later subscribed to by the Netherlands, Switzerland, and Belgium. The Tripartite Agreement, as it is commonly called, substitutes coöperation for competition in exchange management. It is to be credited with some minimizing of exchange fluctuations, and on several occasions it made it possible for France to devalue the franc with less disturbance than would otherwise have been caused. The weakness of the agreement lies in the fact that any one of the signers can withdraw from it at very short notice.

American exchange is definitely a controlled exchange. The power of the Executive to alter the weight ascribed to the gold coin, the authority of the Secretary of the Treasury to monopolize dealings in gold, and the operations of the stabilization fund all point to this, and there is the further power vested in the President to bring all foreign exchange transactions and all international transfers of credit under governmental regulation and control in the

event of any national emergency as declared by him. It is true today that the daily routine of foreign transfers is not greatly affected by executive control, but none the less the authority exists to put into effect a much more rigorous control.

Conclusions. Study of recent developments in foreign exchange seems to show that the present trend is in the direction of a greater degree of control over the exchanges rather than a return to the free exchange market of former times. Many students of monetary theory believe this is a desirable trend, and certainly those who advocate a managed currency divorced from all connection with gold are forced to a belief in a controlled foreign exchange market. The weight of the evidence would seem however to indicate, as we have previously concluded,[1] that the virtues of an international gold standard outweigh its defects. Although the improbability of securing in the immediate future a general return of nations to the gold standard may be granted, it would appear that international coöperation to select appropriate gold parities and to bring about a more equitable distribution of the gold reserves of the world offers one of the most profitable approaches to a return of world prosperity. That stable exchanges and the removal of exchange restrictions would lead to an increase in world trade and in world financial dealing seems obvious. It might indeed be true that certain of the less drastic methods of control could be retained with profit, or that international coöperation might at the start find it wise to revise parities if the ones originally chosen were shown to be exerting an unwholesome influence on the balance of payments. But the strangulation of world trade by exchange controls, the futile competition in exchange depreciation in order to force exports, the singularly vicious attempt to balance the exports and imports between every pair of countries, and other such practices cannot continue in a world that demands the advantages of international division of labor.

EXERCISES

1. Suppose that as a result of a reduction in weight of the American monetary unit its value in terms of foreign currencies declines by 40 per cent. What is the effect of the depreciation of the dollar on each of the following transactions?

[1] See Chapter XXVIII.

(*a*) An export of American automobiles valued at $100,000
(*b*) The import of French wine valued at 100,000 francs
(*c*) Repayment by a foreign corporation of $10,000 due to an American bondholder
(*d*) The expenses of American tourists abroad
(*e*) War debt payments to the United States

2 (*a*) How would the rate of exchange with France be affected by transfers of investment funds by Frenchmen to this country?
(*b*) How would this development affect the export-import trade with France?
(*c*) How might the American stabilization fund be employed to offset the transfers of investment funds to this country?
(*d*) How might the stabilization fund be used to offset withdrawals of investment funds from this country?

XXXVII

INDUSTRIAL NATIONALISM AND THE PROTECTIVE TARIFF

International division of labor. Man's struggle for existence, in a world that yields nothing except in response to effort and yields abundantly only to intelligent and skilful effort, inevitably develops division of labor. That division of labor between individuals promotes individual well-being in the economic sense has been fully demonstrated, by experience as well as logic. Though perhaps not quite so self-evident, the advantages to be gained by division of labor between different regions or nations are of the same order and equally demonstrable. This subject has already been developed in connection with our study of interregional trade,[1] where it was made clear that any nation or region obtains for its people the highest possible standard of living (other things being equal) when it devotes its industrial efforts to those lines of production in which it has a comparative advantage and depends for other products upon the industrial efforts of other parts of the world.

National self-sufficiency. In spite of these clear truths, the peoples of the world have never accepted fully as a matter of practical policy the corollary of freedom of international trade. Considerations of various sorts have led to governmental restrictions, more or less severe, upon the trade between nations Sanitary and moral considerations demand restrictions upon the import of certain products, such as narcotic drugs. Tariff duties have been laid upon imports and exports in order to bring in public revenue. National preparedness for war is sometimes believed to require the fostering of certain industries within the nation, even though this may be accomplished only by placing obstacles, such as tariffs, embargoes, and quotas, upon the import of foreign-made products. In recent times, as we have seen,[2] many nations have believed that

[1] See Chapter XXXIII. [2] See Chapter XXXVI.

the protection or control of their monetary systems and the stabilization of their foreign exchanges required them to place restrictions upon imports. But most potent of all the influences leading to restriction of the freedom of international trade has been a popular belief, widespread and persistent, that by restricting its foreign trade, especially by checking imports, a nation may enhance its economic well-being in terms of greater wealth and a higher standard of living. It is this motive, more than any other, that has occasioned the protective tariffs by which many nations have seriously curtailed their imports of foreign-made goods.

In consequence of all these considerations — and especially the last — there has developed an ideal of national economic organization which is the antithesis of international division of labor and which may be called "national self-sufficiency" or "industrial nationalism." This ideal requires that a nation be, so far as possible, free from reliance for important products upon the industries of other peoples. For its accomplishment there is required obviously severe restriction of imports from abroad; the extreme limit would be prohibition of all imports. Few would actually urge this limit, but a tendency to regard all imports as a menace to national well-being is evident. While in strict logic the ideal of nationalism does not necessarily imply a forcing of exports, the fact is that those who hold to this ideal generally regard exports with a favor no less than their disapproval of imports. Imports are bad; exports are good; the former impoverish, the latter enrich, the nation. Here is the origin, springing from the Mercantilists, of our technical terms "favorable" and "unfavorable" balance of trade.

This ideal of industrial nationalism has been entertained, more or less widely, ever since men began to think of economic problems in modern terms. It was a cardinal principle in the philosophy of the Mercantilists, and its accomplishment was consciously sought in the practical policies of the leading nations of Europe in the seventeenth and eighteenth centuries. It was attacked with devastating effect by Adam Smith. It later acquired a new lease on life through the teachings of the German economists of the so-called "historical school." It is an article of faith of the protectionist movement which has flourished in many nations for many generations. At the

moment, the ideal of industrial nationalism has gained a hold upon popular thinking stronger perhaps than ever before since the days of the Mercantilists. It is closely interwoven with commonly held doctrines of the causes and cure of depression and is apt to be an essential part of current conceptions of "the planned economy." It must finally not be forgotten that national self-sufficiency has ever been urged by many whose own self-interest would be promoted either by the restriction of imports generally or — what is more to the point — of the particular product which they are making in competition with foreign producers.

Effects of the protective tariff. For the accomplishment of the ends of national self-sufficiency a variety of means is at hand. Embargoes, limitation of permitted imports by national quotas, restrictions upon foreign exchange, depreciation of the national monetary unit — all these are devices which have been employed, never perhaps so widely or vigorously as at the present time. But the chief instrument for the restriction of imports is the protective tariff, and the whole subject of industrial nationalism may be elucidated in large measure by analysis of the problem of protectionism.

For our excursion into this subject the principles of international trade and the principles of the shifting and incidence of import duties, as developed in earlier chapters,[1] furnish the starting point. We know that, in the absence of restrictions upon its foreign trade, the people of each nation tend to buy goods in the cheapest market and sell in the dearest market and that this leads them to devote their productive resources and energies to those lines in which their nation has comparative advantage, as evidenced by the ability to produce at comparatively low cost. At the same time they rely upon other nations for those goods in whose production they are at comparative disadvantage. Thus develop international trade and international division of labor.

International trade may be checked by means of duties on imports.[2] As we have seen, a duty high enough to be protective may

[1] Chapters XXXIII, XXXV, and XXXI

[2] Since they seldom enter into the controversy — never in the United States — no consideration need be given in this discussion to duties on exports.

diminish or entirely stop the import of any chosen commodity. Carried to its extreme limit, the protective tariff might be used to prevent virtually all imports. Even if applied only to those goods which are capable of production in the home country, the result would not be much different, since barring only those commodities whose materials are not to be found within the territory, there is virtually no commodity which could not conceivably be produced if the foreign product were kept out. Doubtless such exotic products as tea and coffee would be produced in the United States -- of inferior quality and at fabulous cost to be sure – if a high enough tariff excluded the imported articles. Whether applied moderately or to the limit, the protective tariff has the effect of reducing imports and building up corresponding industries at home. It thus restricts international trade, promotes "diversity of industry" in the home country, and diminishes reliance upon the goods of other nations. National self-sufficiency gains at the expense of international division of labor.

As to the truth of the foregoing statements no question can be entertained by anyone who is familiar with the elementary principles of economics. The controversy over protection starts from this point. Granting the respective effects of freedom of trade and protection as we have stated them, the "protectionist" asserts that it is to the economic advantage of a nation thus to increase diversity of industry and national self-sufficiency; the "free-trader"[1] says that this means an economic loss and that the nation gains most when it enjoys all the advantages of international division of labor. Thus the issue is sharply drawn.

Protection and national self-sufficiency: The case for protection. The advantages of trade, as already developed, are so evident and so important that we are justified in placing the burden of proof upon the protectionist, who claims that there is an economic gain from putting artificial restrictions upon a nation's foreign trade. Let us proceed then to examine the principal arguments advanced

[1] The reader should perhaps be reminded that "free trade" as the term is used, does not mean no tariff, it means no protection. The free-trader has no quarrel with the revenue tariff, he is opposed only to protection. There is no country in the world entirely without a tariff and no appreciable sentiment in favor of the abolition of all tariffs.

in support of the policy of protection as a permanent device for giving a nation a greater measure of self-sufficiency.

1. The favorable balance of trade argument. Some of the most popular arguments for protection are such naive applications of old Mercantilist fallacies that they may be dismissed with nothing more than brief reference to the established principles of trade. Such is the argument that a nation gains in its foreign trade only from its exports, that imports are a national loss, and that the tariff is therefore beneficial through its effect in checking imports. We have already learned that, so far from exports being superior, it is imports which are the real goal of international trade; we have learned also that a country's balance of payments must in the long run be in substantial equilibrium and that therefore anything which reduces imports must eventually also reduce exports. It is neither advantageous nor possible by means of a tariff to make imports decline relatively to exports.

2. The gold argument. Similarly naive is the argument that a favorable balance of trade, created by the tariff, causes gold to flow into the country and is therefore advantageous. This is simply a corollary of the preceding argument and rests upon other Mercantilist fallacies with which the reader is acquainted. We know that it is not the balance of trade (*i.e.*, merchandise exports and imports), but the balance of payments (including all debits and credits) which determines the inflow or outflow of gold; and we are also aware that it is neither possible nor advantageous for a country to pile up gold as the result of artificial restrictions upon its foreign trade.

But there are those who are not content with this summary disposal of cherished beliefs. Is it not true, they say, that the rule for amassing wealth is to sell much and buy little? Is it not thus that the individual business man accumulates money from the profits of his business? And can we avoid the conclusion that the same rule must hold for national gain from trade? At the risk of adding proof of the obvious, we may give some further attention to these Mercantilist notions.

The wealth of the people of the United States in 1925 was estimated at about 355 billion dollars. At that time the nation's total

stock of gold coin and bullion was about $4\frac{1}{2}$ billions, or about $1\frac{1}{3}$ per cent of the national wealth. In 1880 the national wealth was, according to estimates of the United States Census, about 44 billions. The stock of gold coin and bullion was then 352 millions. In this period of 45 years therefore the national wealth was increased by 311 billions, while slightly over 4 billions were added to the stock of gold. It is certain enough that this great gain in national wealth was not due to accumulation of gold from foreign trade. The growth of a people's wealth is not principally in the form of gold; it consists in the increased value of its lands and houses, its factories and railroads and ships, its homes, and countless other forms of wealth vastly more important than its stock of gold. And the gain from its foreign trade is not in the form of gold but in the useful commodities imported.

But what of the analogy of the business man? The answer is that neither does he accumulate wealth in the form of money or by selling more than he buys. If at the end of a profitable year in which a man has made gains of $50,000, we should ask him to show us his profits in concrete form, he would not point to an increase of his money on hand or his bank deposits. The probability is that these items are no larger than at the beginning of the year. His profits were not made by selling more than he bought. As he received money from the sale of goods, he spent it — partly for current consumption, partly for additions to his factory and machinery, for lands, for improvements to his home, for a new automobile, for government bonds and corporation securities. Thus he will show concrete evidence of his year's accumulation So it is with the nation. A nation, like an individual, grows wealthy, not by selling more than it buys, but by producing more than it consumes

3. The added industry argument. It is argued that the new industries which admittedly may result from protection are an addition to the industrial resources of the nation and that therefore the tariff has increased the national wealth. Let the reader recall that the tariff causes the establishment of a new industry by making it too costly to buy a foreign product previously imported. He will further recall that this diminution of our imports must be balanced eventually by a corresponding decline of our exports. The balance

of payments must maintain its equilibrium; if we will not buy of foreigners, they cannot buy of us. The result is that, while a new industry is arising, some old industry finds its foreign market declining and is forced to reduce its operations accordingly. The new industry is thus not an addition to the nation's industry but merely a diversion of capital and labor from one field to another.

It is true that, when a country is growing in population and enterprise, the diversion of industry induced by protection may take place without causing an absolute decline in any field. What the tariff does in this case is to direct into the protected industry new capital and labor which would otherwise have gone elsewhere. The development of some unprotected industry is stopped or diminished.

In either case, whether capital and labor already employed are shifted or new capital and labor are directed into the protected field, there is clearly no gain in total industry. Moreover this diversion of industry is from the more profitable into the less profitable fields. In the absence of restrictive legislation, the nation's capital and labor are employed in those lines in which it has the greatest comparative advantage, and it is into these same lines that new capital and labor seeking employment in a growing country will come. The protective tariff prevents this natural disposition, forces industry into less advantageous lines, and so causes a net economic loss.

4. The wages argument. This argument appears in two forms. (1) It is asserted that the protective tariff, by increasing the total amount of industry in the country, adds to the demand for labor and so tends to increase wages. Protection is therefore of advantage, not merely to the entrepreneur class, but to all laborers, and so to at least the great majority of the people. It is evident that this argument rests for its foundation on the added industry argument and falls with it. The mere diversion of industry, particularly when it is from more advantageous to less advantageous lines, can at best do nothing to increase the demand for labor or to raise wages.

(2) The wages argument is not always put so crudely as this. By American protectionists it is more often argued that, in the

absence of the tariff to keep out cheap foreign goods, American manufacturers could not pay their present high scales of wages but would be compelled to lower wages or perhaps to go out of business entirely. Thus the tariff is in the interest of the workingman. This argument overlooks the fact that many American industries seem able to get on well enough without protection. These are the industries naturally fitted to prosper in the American environment, the industries in which America has a comparative advantage. Without protection, they are nevertheless able to pay the prevailing high level of American wages; they pay as high wages as the protected industries. If any industry is really unable to pay the prevailing rate of American wages it is because that industry is not adapted to our conditions. Its presence is due to the artificial stimulus of the tariff, and it exists at the expense of some more favorable industry which could pay high wages without tariff assistance. Removal of the tariff might cause the artificially favored industry to disappear, but its loss would be offset by the gain in the naturally favored industry, and there would be no decline in wages.

Wages and labor cost. There is much loose talk about wages in this connection. Protectionists argue as though, simply because wages are higher here than elsewhere, it must cost more to produce anything in America. They express alarm at the prospect of competition with the products of "foreign pauper labor." Now it is a well-known fact that the wages of farm laborers in the Mississippi valley are higher than anywhere else in the world; yet there are few regions where the cost of producing wheat is so low. In numerous fields of manufacture goods are produced more cheaply in America than anywhere else in the world, in spite of high wages. The explanation of low production costs in such cases is of course the natural advantages of soil, climate, and mineral resources, the unsurpassed use of labor-saving machinery, and the extraordinary skill with which certain American industries have been organized. Under such conditions employers are able to pay high wages for high grade labor, and that without making the cost of production high.

American industry could not function on its present level of high

efficiency if it had to do with only the foreign pauper labor to which the protectionist argument refers. Few American employers would be willing to exchange their present personnel of skilled labor for a complement of Chinese coolies even at the Chinese level of wages. The student will recognize that the countries whose competition our manufacturers most fear are not India, China, or Africa, but England, Germany, and France. With few exceptions, our most effective competitors are the countries of comparatively high wages. And the competition which European manufacturers fear includes America, the country of the highest wages of them all. In fact it is notable that among nations industrial efficiency goes with high wages, not with low wages.

In international competition, as in industrial competition everywhere, the controlling factor is not high or low wages, but high or low *labor cost of production*. Wages, as such, tell us nothing. It is only when wages are divided by quantity produced, to give the labor cost per unit of product, that we have the key to the problem of relative advantage in international competition. Thus it is that high relative wages when accompanied by still higher relative output give low labor cost. This is the condition in the great body of industries which are naturally adapted to American conditions. Here high wages are paid without the aid of protection. In general these are industries which either are favored by peculiar advantages of soil or climate or natural resources, such as cotton growing, wheat growing, copper mining, etc., or have the advantage of skillful use of labor-saving machinery, such as the manufacture of automobiles, sewing machines, farm machinery, etc. Industries not so favored by nature or which must lean heavily upon hand labor do not flourish here. To them the high level of American wages is a real handicap, which can be overcome only by the protective tariff removing the competition of the cheaper foreign product. Protection is required therefore, not to maintain the high level of American wages, but to make possible the payment of American wages in industries not suited to American conditions.

Historically, high wages in America antedated protection, and it was these very high wages that led to the demand for protection in lines which otherwise could not attract labor from those fields in

which America offered the greatest rewards. In this sense it is correct to say that, so far from the tariff being the cause of high wages, high wages were the cause of the tariff.

5. Effect of withdrawing protection. The wages argument is without doubt urged in good faith by many American manufacturers who honestly believe that protection is necessary, not only to their own profits, but also to the well-being of American labor. Being engaged in lines of industry which possibly could not exist without tariff protection, they realize that without protection they could not continue giving employment to their labor forces at the present scale of wages. They visualize the distress which would follow a material cut in wages or a shut-down of industrial plants. They fail to see that discontinuance of the artificially fostered industry would eventually open up other fields in which these laborers would find employment at American wages.

It is not to be denied that the shifting would impose inconvenience and loss upon the laborers and still more upon the capitalists. Industry that has become adapted to the artificial condition of protection cannot adjust itself to freedom of trade instantly or without some loss to all concerned. This is no argument for retaining permanently the artificial condition But it does dictate caution and consideration in making the change. Protection has long been an established American institution; entrepreneurs have directed capital and labor into certain fields of industry relying upon the government's policy of protection. They and their employees have the right to ask that a change in policy be made gradually and with every possible care to minimize the unavoidable losses of adjustment.

The fact is that American industry would not be wrecked by a careful gradual removal of protection. Some few industries would doubtless succumb. In most cases certain grades or lines would be discontinued while others advanced, the sort of adjustment which normally goes on in every progressive industry. For industry as a whole the final effect would without question be beneficial. American tariff history shows protection surprisingly less potent, either for good or for evil, than both its friends and its enemies would have us suppose. The course of American industry has not been vitally

influenced by its imposition, nor would it now be revolutionized by its withdrawal.

6. Alleged saving in transportation costs. As an argument in favor of protection, attention is sometimes called to the waste involved in sending raw materials to a foreign country only to be returned to the country of origin in the form of manufactured products. An example is American cotton, much of which goes abroad to English cotton mills, and some of which finds its way back to America in fine British cotton fabrics. Here is said to be a wasteful double cost of transportation, which might be avoided by a tariff which would keep out the British product and compel its manufacture in America.

The fallacy here is obvious. There is no loss in the operation in question; otherwise it would not occur. Transportation costs are no worse than other costs. The double transportation of cotton across the Atlantic is undertaken because that is the way to get the product for the American consumer at the least total cost. Forbidding this method by a tariff would save this transportation cost, to be sure, but only at the expense of a still greater cost of manufacture. The *reductio ad absurdum* of this argument appears if we imagine ourselves trying to apply it universally. Practically all domestic industry involves such double costs of transportation. Wheat is raised on Nebraska farms, shipped to Minneapolis, ground into flour, and returned to the Nebraska farms. Iron and copper and coal are brought from Minnesota and West Virginia to New England, and the finished steel and brass goods are shipped back to these states. The sole reason for all this double transportation is to minimize total costs of production. Its avoidance would be possible only at the cost of abandoning territorial division of labor and putting the human race back in the Middle Ages economically.

7. Meeting the tariffs of other nations. There are those who are ready to admit the advantages of universal freedom of trade between nations but who yet feel that, in a world in which protective tariffs are common, no one nation can afford to permit free entry of foreign goods. Nor can one nation safely remove its tariff barriers unless other nations do the same. This misgiving overlooks the

experience of Great Britain, which, during half a century in which she led the world in an amazing industrial development, maintained freedom of trade regardless of the tariffs imposed against her products by other nations. It overlooks also the fundamental principle. It is true that our own industries and our economic welfare are injured by the tariff barriers of other nations, which deprive us of some measure of the benefits of international trade. But the whole conclusion of our analysis has shown that our national welfare is injured by the curtailment of our foreign trade occasioned by our own tariff. Freeing our trade from the shackles both of our own tariff and the tariffs of foreign nations would be doubly advantageous. But if we cannot secure the removal of foreign tariffs, that is no reason why we should forego the gain to be derived from removal of our own tariff.

The trouble with reciprocity agreements, tariff treaties, and the like is that they are too often conceived in the spirit of Mercantilism, which regards exports as a national blessing and imports as a national curse. There follows a sort of inverted horse-trading negotiation, in which each nation seeks to give the other the utmost of its products while taking as little as possible in return. Once the notion is reversed and imports are recognized as the true goal of foreign trade, with exports desirable only as the means whereby imports may be obtained, the whole spirit of mutual tariff concessions is changed. No nation need grudge other nations whatever advantage they may derive from relaxation of its own tariff. If concessions in foreign tariffs may be obtained in return for reduction of our own tariff, that is all to the good, but it does not follow that a proposal to relinquish our own trade obstructions should be given up merely because other nations do not see fit to do likewise.

8. Mutual dependence of nations. Mercantilist notions sometimes run to the extreme belief that whatever injures the industry and trade of another nation must therefore be beneficial to us. It follows that one of the chief services of our tariff policy should be to injure the economy of foreign nations by excluding their goods from our markets. That can of course be done. We have seen that in that case the principal injury is to ourselves. It now remains to point out that we suffer also from whatever injury our tariff policy

may inflict upon other nations. The whole world is bound together by the universal principle of division of labor. Just in proportion as the industries of other nations are efficient and productive will they be able to serve our needs by enabling us to obtain through exchange certain products from them at less cost than if we had to make them ourselves. In general each nation stands to gain economically, not from the adversity, but from the prosperity of other nations.

Considerations against protection. Having placed the burden of proof on the protectionist and found his arguments lacking in probative force, the free-trader might conceivably rest his case here. There are however some important positive considerations on the free trade side of the argument which may serve to throw further light upon the problem.

What is the territorial unit for protection? It is an interesting characteristic of the foregoing arguments for protection that each assumes that the tariff barrier will follow national boundary lines. Yet the arguments themselves are purely economic; they contain nothing to show that protection would not be exactly as beneficial if directed against a competing region within the same nation. We are tempted to ask why the New England textile mills should be denied protection against the rising industry of the South, and *vice versa*, or why the automobile industry in various states should not have protection against Detroit. Historically many of the first tariffs were imposed by towns or other minute political subdivisions. They proved such a burden on trade and such a drag on economic progress that they were swept away. In like manner it came to be recognized that the tariff systems of some score or more little states in central Europe were proving fatal to the economic development of that region, and the German *Zollverein* — customs union — swept them all away and opened the path to a later political union and to the marvellous industrial development of the German Empire. Originally the American states had each its own tariff system directed against the others. They had to give way in the interest of the economic welfare of the whole region. No one today would seriously propose to retrace any of these steps already taken in the direction of freedom of interregional trade. No argument exists to

show that still further progress would not follow the removal of the remaining barriers to the trade between nations.

Benefits of foreign competition. It is a fallacy to regard foreign competition as especially heinous. Manufacturers naturally dislike competition. They seek to avoid it whenever they can. But competition is in the interest of the consumers, for whom industry in the last analysis exists. And the competition of foreign producers is just as wholesome as that which is nearer at hand. If the home manufacturer cannot meet the foreign competition, then the consumers are entitled to the less costly product. If the home manufacturer, spurred on by the foreign competition, does succeed in meeting it and holding his market, then again the consumers are benefited. Moreover the nation gains from the increased efficiency of an industry which might otherwise have lagged behind. Instances are common enough of American industries which, if pushed to do their utmost, could have competed on even terms with the best of the foreign producers, but which, sheltered by the American tariff wall, have grown lax and fallen behind the industrial procession.

Conclusion. Analysis of the arguments for and against protection thus appears to demonstrate that industrial self-sufficiency secured by means of the protective tariff, so far from enhancing the national welfare, is a cause of economic loss. No gain can be shown to offset the demonstrated advantages of international division of labor. Protection is an economic burden, a cause of loss to the people of the nation which employs it. The economic advantage is all on the side of freedom of international trade.

It is probably not necessary to remind the reader that these conclusions do not imply that any nation — least of all the United States — could instantaneously abolish its entire tariff system without some unhappy consequences. Tariff systems have grown up during a long period, and the industry of each nation has adapted itself to the tariff situation as it has been and is. In America a system of very high tariff duties upon most of the important products of manufacture and upon some raw materials has been in effect for three-quarters of a century. There have in this period been variations in the level of protection, but never anything approaching free

trade. American industry has grown up in this environment, to which it has had constantly to adapt itself. To change, suddenly and without notice, this condition by going over to the free entry of all foreign-made goods would throw industry into confusion and do serious injury to the economic welfare of the nation. From this proposition — which must be well-nigh self-evident — the conclusion is, not that free trade is not the ideal or that protection must be retained, but rather that the change from a condition of protectionism to the healthier condition of free trade must be made gradually and with due notice, to the end that the costs of readjustment to industry and the inconvenience to the people be made as small as possible.

Temporary protection. Thus far we have been considering the question of protectionism from the viewpoint of the relative merits of international division of labor and national self-sufficiency, the protective tariff serving as a permanent device for promoting the latter ideal. In contrast we must now observe the possibility that the tariff may be called upon from time to time to serve certain temporary objectives not involving any permanent artificial interference with the freedom of international trade. The classical example of this application of protection is the use of the protective tariff on behalf of young industries.

Infant industry protection. Use of the protective tariff has been frequently urged and occasionally employed as a means to facilitate and hasten transition from the agricultural to the industrial stage of national economy. A nation on the threshold of this change will be taking the first steps toward establishing numerous manufacturing industries. In these first steps manufacturing will have to contend with the competition of imported goods coming from older nations whose industries are already firmly established. They have the machinery, the technical knowledge, the skilled workers; they are directed by experienced entrepreneurs. None of these essentials has as yet been attained by the young industries of the first country, and the competition of foreign manufacturers is felt as a severe handicap.

Under these circumstances the tariff may be used to give relief from foreign competition until such time as the young industries

have become firmly established. This line of argument is known as "the infant industry argument" for protection. When logically used it rests upon the following assumptions: (1) the argument applies only to industries which are suited to the national conditions, (2) the obstacles to the establishment of such industries are temporary only, (3) the tariff protection need be temporary only; after the removal of the tariff the industries will be able to stand alone. When these conditions are fulfilled and when it is further proved that the nation will actually benefit from the establishment of the industries in question more rapidly than they would come in the natural order of events, the infant industry argument is undoubtedly sound.

Conditions in the United States at the beginning of the nineteenth century were fairly typical of those here assumed, and the first tariff acts were to some extent at least justified by this argument. Even more than the tariff, the international conflicts of the period from 1807 to 1815, by interfering with foreign trade, gave protection to the young American industries. But the conditions to which the infant industry argument was applicable had generally passed with the first third of the century. This argument is little heard of today. It must be noted also that it has not often been urged in good faith; having used it to obtain tariff favors, its advocates have later shifted to other arguments to prevent removal of protection. Most protectionists today advocate permanent protection.

Protection against "dumping." From time to time producers of some nation undertake to unload a surplus product upon foreign markets at prices lower than are charged in their own domestic market, sometimes even lower than cost of production. When goods are thus exported at less than cost the process is called *dumping*. Dumping may be resorted to by some particular firm whose position sufficiently approaches monopoly, or by a combination of producers acting in concert for this purpose, for example the closely organized chemical industry of Germany. Often such export trade is encouraged or actually subsidized by the government. For a long time it was the practice of the state-owned railroads of Germany to grant special low freight rates to certain great kartels on goods

destined for export. During the past decade dumping has become more and more a governmental policy, especially of certain authoritarian nations, employed in connection with their efforts to control their foreign trade and foreign exchanges and protect their stocks of monetary gold. Similar results are caused when nations, as has been quite common in recent years, deliberately devalue their currencies in order to take advantage of the temporary stimulation to export trade which tends to follow.

The producers of any nation whose markets are invaded by such dumping of foreign goods are subjected to competition of special severity. Were this a stable and presumably permanent condition, it would be the part of wisdom for the nation concerned to accept the situation, congratulate itself that its people were able to obtain some needed products thus cheaply, and let its own capital and labor transfer to more advantageous lines. But dumping, from its very nature, is fickle and unstable. Such temporary competition may work serious injury to a nation's industries, only to leave its people poorly supplied when the foreign producers return to a more normal price policy. The reader will recognize the similarity between foreign dumping and certain practices of monopolistically inclined enterprises in the domestic market, practices which have earned the title of "unfair competition."

It is only natural that a nation should seek to protect its own producers against dumping. The protective tariff is the obvious means. Wisely employed, the tariff may serve to offset the price cutting of foreign producers, with the result that the foreign goods enter the domestic market at prices not less than their normal costs of production. This policy involves frequent adjusting of tariff rates to meet changing policies of foreign producers, and it requires skillful handling. There is moreover always the danger that selfish interests will use the dumping bogie to secure excessive or permanent protection. Suitably employed however protection against temporary dumping is in the interest of the whole people of the nation concerned. Protective tariffs and other measures against dumping are at present being employed by an increasing number of nations.

Protection on account of military necessity. Protection may be urged from an entirely different standpoint. The tariff is necessary,

it is said, in order that the nation may be assured of a supply of certain products vital to its safety in time of war. If it be proved (1) that a certain industry is thus necessary and (2) that this industry will not exist unless relieved of foreign competition, then it follows unquestionably that protection is advantageous. Whether these two assumptions are actually fulfilled is a matter of fact to be determined in each case; it is doubtful whether, in a country of the size and variety of resources and industries of the United States, there is much room for this line of argument.

Without doubt the extraordinary development in recent years of protectionism among the nations of Europe has been in large measure based upon military considerations. These nations are preparing for war and are seeking the utmost of national self-sufficiency in order that in the next war they will not be weakened by enemy blockades as was Germany for example in the World War. Thus we are seeing tariff rates of unprecedented height, embargoes, quotas, and the other devices which have marked the nationalistic hysteria of the past decade or two.

It should be noted that the military necessity argument for protection is not based upon economic grounds. There is no false assumption that the people will be richer on account of the tariff. It is admitted that they will be poorer, through having to pay more for some important commodity than they need have. But so are they made poorer by the taxes which defray the cost of building battleships. The justification is the same in either case.

Protection by bounties. Occasionally the object of protection is attained, not by a duty on the foreign product, but by a bounty paid by the government to the domestic producers. The United States for example has on occasions sought to encourage the growing of sugar by means of bounties. As to their economic significance the bounty and the protective tariff are alike, with one important exception. The cost of protection by tariff is imposed upon the consumers in proportion to their use of the goods in question. The cost of the bounty rests upon the whole group of taxpayers. Since protection can be justified, if at all, only on the grounds of public advantage, the apportionment of the cost under the system of bounties is more equitable than that of the protective tariff. Never-

theless bounties are very rarely used, probably because it is here so evident that protection costs something, that a burden is placed on the whole people for the benefit of a favored few.

Ship subsidies. Ship subsidies are bounties paid to ship owners or operators in order to encourage the development of national shipping -- the "merchant marine." The economic nature of ship subsidies is precisely the same as of protective tariffs and bounties. The justification of subsidies to shipping must be found, if at all, in the military necessity argument.

Preferential trade and tariff arrangements. At the same time that the recent wave of extreme nationalism has led most nations to curtail severely their trade with other nations, there has developed a contrary tendency to enlarge the area of comparative freedom of trade by means of special tariff agreements. The British Empire furnishes a conspicuous example. In 1932 Great Britain took definite action terminating her traditional policy of free trade, and in the same year a scheme of imperial preferences was set up by agreement among Great Britain and the various self-governing dominions. Thus was the British Empire bound together in an area of tariff preference as against the rest of the world. France, without going so far toward a formal organization, has made preferential tariff arrangements with her colonies. Germany has been seeking to build up a preferential tariff group, including nations under her special influence in central Europe as well as certain South American countries. Japan is quite evidently seeking to dominate a great Eastern tariff area including Manchukuo and China.

Reciprocal trade agreements of the United States. A somewhat different line has been followed by the United States. The trade agreement act of 1934 (continued by act of 1937) authorized the President to enter into foreign trade agreements with other nations, in connection with which he was empowered to make reductions up to 50 per cent in existing tariff rates upon certain imports from a given nation in return for corresponding concessions in that nation's tariff in favor of American exports. The Secretary of State, Cordell Hull, has acted with great vigor in negotiating such reciprocal tariff arrangements. By the fall of 1939, 21 such agreements had been concluded, the most important, and possibly the most

difficult to arrange, being the agreement with Great Britain and a new agreement with Canada, which were signed on November 17, 1938, to go into effect on January 1, 1939.

These agreements are reciprocal, but not preferential like those of most other nations, since the United States still lives up to the liberal "most favored nation" principle, in accordance with which any reduction granted to one country is "generalized" to all nations with which the United States has most favored nation treaties. This does not mean that there is no advantage in being the party to a trade agreement with this country, since the agreements are based on the "chief supplier" principle; *i.e.*, the United States makes tariff reductions only on articles coming exclusively or chiefly from the nation which is the other party to the agreement.

It would be too much to suppose that this reciprocal agreement policy is making any real breach in the traditional wall of American protectionism. Indeed care has been taken not to permit competition that would seriously affect any American industry. This is accomplished both by the choice of articles for tariff concessions and by limiting the amount of import; for example, in the agreement with Czechoslovakia her producers were permitted to send shoes at the reduced rate, but only up to 1.25 per cent of the average annual production of the United States. Nevertheless in a world that seems rapidly to be throwing away important economic benefits through an exaggerated nationalism, the present attitude of the United States is furnishing a powerful influence on the more liberal side. This comes with special significance from the nation which for a century has been the chief exponent of the policy of high tariff protection.

SUGGESTIONS FOR FURTHER READING ON PART VI, INTERREGIONAL TRADE AND EXCHANGE

I. Interregional Trade

Taussig, F. W. *International Trade* New York, 1927
Ohlin, B. *Interregional and International Trade*. Harvard University Press, 1933
Haberler, G von *The Theory of International Trade* New York, 1936
Killough, H. B *International Trade*. New York, 1938
Ellsworth, P. T. *International Economics* New York, 1938

de Haas, J. A. *The Practice of Foreign Trade.* New York, 1935
Griffin, C. E. *Principles of Foreign Trade.* New York, 1934
Horn, P. V. *International Trade.* New York, 1935
Pratt, E. E. *International Trade in Staple Commodities.* New York, 1928
Sinclair, H. M. *The Principles of International Trade.* New York, 1932
Zimmerman, E. W. *World Resources and Industries.* New York, 1933

II. Foreign Exchange

Escher, F. *Foreign Exchange Explained.* New York, 1917
Furniss, E. S. *Foreign Exchange.* Boston, 1922
Whitaker, A. C. *Foreign Exchange.* New York, 1933
Gayer, A. D. *Monetary Policy and Economic Stabilization.* New York, 1935
Einzig, P. *Exchange Control.* London, 1934
—— *World Finance, 1914–1935.* New York, 1935
—— *Exchange Clearing Systems.* New York, 1935
Hall, N. F *The Exchange Equalization Account* London, 1935
Harris, S. E. *Exchange Depreciation.* Cambridge, 1936
Madden, J T., and Nadler, M. *The International Money Markets.* New York, 1935
Salter, Sir Arthur. *China and Silver.* New York, 1934
Westerfield, R B. *Our Silver Debacle.* New York, 1936
National Industrial Conference Board. *The New Monetary System of the United States.* New York, 1934
The Future of Monetary Policy. A report on International Monetary Problems by a Group of the Royal Institute of International Affairs. London, 1935

III. Commercial Policy and Protective Tariff

Williams, B. H. *Economic Foreign Policy of the United States.* New York, 1929
International Economic Relations. The Report of the Commission of Inquiry into National Policy in International Economic Relations. University of Minnesota Press, 1934
Taussig, F. W. *Tariff History of the United States.* Eighth edition. New York, 1931
——, and White, H. D. *Some Aspects of the Tariff Question.* Harvard University Press, 1931
Beveridge, Sir W. H. *Tariffs, the Case Examined.* London, 1932
Jones, J. M., Jr. *Tariff Retaliation.* University of Pennsylvania Press, 1934
Taylor, A. E. *The New Deal and Foreign Trade.* New York, 1935
Tasca, H. J. *The Reciprocal Trade Policy of the United States.* Philadelphia, 1938
Smith, J. G. *Economic Planning and the Tariff* Princeton University Press, 1934
Beard, C. A., and Smith, H. E. *The Idea of National Interest.* New York, 1934

BEARD, C. A., and SMITH, H. E. *The Open Door at Home.* New York, 1934
MARTIN, R. F. *International Raw Commodity Price Control.* National Industrial Conference Board. New York, 1937
GRAHAM, F. D. *Protective Tariffs.* New York, 1934
ASHLEY, P. *Modern Tariff History — Germany, United States, France.* New York, 1911
DELLE-DONNE, O. *European Tariff Policies since the World War.* New York, 1928
ASHLEY, W. J. *The Tariff Problem.* London, 1911
HIRST, F. W. *Safeguarding and Protection in Great Britain and the United States.* New York, 1927
PATTEN, S. N. *The Economic Basis of Protection.* Philadelphia, 1890
DIXWELL, G. B. *The Premises of Free Trade Examined.* Cambridge, Eng., 1881
PIERCE, F. *The Tariff and the Trusts.* New York, 1913
STANWOOD, E. *American Tariff Controversies in the Nineteenth Century.* New York, 1903
RICHARDSON, J. H. *British Economic Foreign Policy.* New York, 1937
PEEL, G. *The Economic Policy of France.* New York, 1937
STALEY, E. *War and the Private Investor.* New York, 1935
—— *Raw Materials in Peace and War.* New York, 1937

PART VII

THE INDUSTRIAL ORGANIZATION

XXXVIII

RISK, INSURANCE, AND SPECULATION

Futurity, human knowledge, and risk. The most important attribute of the present moment is that it is the dividing point between past and future. The past is gone; nothing can be done to change it. The future is before us, and it is for the sake of future effects — satisfactions to be secured and pains to be avoided — that all economic activity is put forth.

Now while personal experience and recorded history tell us with more or less certainty of past events, the future is notoriously uncertain. Man cannot say that he is absolutely certain of any future event, though the degree of uncertainty varies all the way from the almost complete certainty that the sun will rise tomorrow to the almost complete uncertainty as to what the direction of the wind will be a year from today. The very nature of the universe in which man lives compels him to bear *risk; i.e., the chance of loss or injury.* Since the uncertainty of future events permeates every phase of human existence, risk is universal. It takes a thousand forms — the uncertainties of death, disease, unemployment, fire, change of price, defective title, rise of costs, change of fashion, strikes, spoliation, piracy, tariff changes, tornado, earthquake, political instability, strength of materials, declaration of war, and so forth without end.

It is evident that the basis of risk is the lack of human knowledge of future events. To an omniscient person, who could foresee with perfect accuracy the future, there would be no risk. To actual human beings risk is always present, because no one is omniscient, and it is small or great according to the degree of our knowledge.

The human attitude toward risk. The instinctive reaction to the unknown is fear, and it is no cause for wonder that primitive man, beset with manifold dangers of whose origin and nature he was ignorant, peopled the universe with malevolent spirits, who might be propitiated by ceremonials and sacrifices. Modern man takes

a calmer and more rational view of risk. He has ceased to regard future uncertainties solely from the standpoint of superstition, though his emancipation is by no means yet complete — witness the common though shame-faced dread of thirteen at table or a wedding on Friday.

In his attitude toward risk, man displays a mixture of emotions. A certain amount of risk now and then is not unwelcome to most of us. Not content with all the innumerable risks which nature provides, we go out of our way to create other hazards. This is *gambling*, the intentional creation, for its own sake, of artificial and unnecessary risk. The man who makes a wager that his football team will win a particular game has created for himself an entirely new risk. Of course certain risks — of chagrin, loss of prestige, etc. — attach in any case to the outcome of the game, but the gambler has added the risk of losing money by the defeat of his team. But while a certain modicum of the gambling spirit is well-nigh universal, men do not relish too much risk, and in general they regard the risks which nature imposes upon them as an evil, to be avoided or reduced wherever possible. In this contest against risk man has a variety of weapons to his hand.

Risk reduced by knowledge. The probability of any event may be defined as the degree of rational belief in the occurrence of that event; every improvement in knowledge raises probabilities toward their highest degree, which is synonymous with certainty. So long as probability is less than certainty, risk is present in human life. When knowledge is sufficient to permit the measurement of probability — a condition in relation to social affairs which implies a broad background of statistical experience — the principles of actuarial reserves and insurance may be applied for the reduction of risks. When knowledge is still less perfect, speculation, forced or voluntary, becomes the only human recourse for combatting the incalculable and uninsurable risks then prevalent.

In the modern era scientific investigation and research have added immensely to our store of knowledge and have thus made prediction of future events increasingly accurate. Man can do little or nothing to prevent the occurrence of storms, floods, earthquakes, and other natural hazards, and our present knowledge of

seismic and meteorological forces is so meagre that earthquakes and weather defy forecasting much in advance. Yet the United States Weather Bureau, in spite of the fact that its forecasts are not altogether dependable even for twenty-four hours in advance, gives information which makes possible avoidance of much loss by provisions taken against forecasted storms. Similarly progress has been made in the business world in analyzing economic conditions, in predicting the train of events, and in attempting to direct the course of business in less risky ways.

Preventive measures. A second method of dealing with risk is to prevent the occurrence of the harmful event. The construction of fireproof buildings, the installation of fire-fighting apparatus, the establishment of safety deposit vaults, the provision of free medical advice and care of insured people, the institution of devices to stabilize the price level — all these are instances of the reduction of risk by preventive measures. New methods of prevention are being constantly devised, and the costs of many of the old methods are being reduced.

This method of dealing with risk is intimately dependent upon the former, for preventive methods cannot be devised until scientific investigation has unearthed the cause of the event. Vaccination against small-pox and inoculation against typhoid fever are two from among the many illustrations provided by medical science.

Combining risks: The law of large numbers. A third method of reducing risk is to combine risks by operating on a large scale. This is based on the principle of scattering risks, of not "putting all one's eggs in one basket," or *the law of large numbers;* namely, that *a high degree of probability may characterize a composite of events of which every individual event is highly uncertain.* Although one cannot know absolutely whether Jones will desert the army, experience will show about what number of men will desert an army of 100,000 men. Of a large group the percentage of men who will die from disease, fail in business, be killed on the railroads, go to college, etc., can be quite closely told in advance. An investor by scattering his funds among a hundred securities makes the risk of loss by default of debtor far less than it would be if his investment were all

in one kind of security. A manufacturer who produces a variety of goods or who markets his product over a broad area with many customers faces less risk of loss of his market than one who clings to one product or to a local market with few buyers. By basing business forecasts and operations on averages of many factors the risk element is reduced and in some cases almost entirely eliminated. The variations of the numerous factors counterbalance one another about the average. In combinations of competitors all those risks arising from ignorance of the probable action of a competitor are eliminated; one of the chief arguments for large-scale industrial combination is the higher stability of business and the consequent lower risks under it.

Reserves. It has long been a commonplace expedient in banking to keep on hand a cash reserve against the contingency of extraordinary sudden demands by the customers of the bank. If funds flowed synchronously to the bank in amounts equal to the withdrawals, no cash reserve would be necessary, but actually the banker can avoid suspension of payments only by carrying some idle cash against the contingency of withdrawals exceeding receipts. The cost of thus eliminating the risk of suspension is measured by the sacrificed interest on the reserve. The expediency of carrying reserves applies to all kinds of business, not to banking alone. To keep his plant going continuously the manufacturer maintains an inventory of raw materials and equipment in excess of the absolute amount which would be required if he could depend upon a regular flow of materials to his plant and continuous service from all his equipment. Likewise he carries a surplus of finished products against the possibility of extraordinary demands from old customers or the demands of new customers. The large corporations find their offerings of securities are much more readily sold if they have had a long record of steady dividends, and to that end they carry dividend reserve funds to make up the deficiencies of lean periods and enable them to declare the regular dividends. In such ways is business stability enhanced and risk reduced by the maintenance of reserves.

It is important to note that by combination the absolute amount of reserve funds may be reduced while still attaining the same

reduction of risk. For illustration suppose it were statistically determined that a tank of 1,000 gallons of water was the minimum of safety to protect a warehouse of 1,000,000 bushels of wheat. If each of ten such warehouses kept its own tank of water, 10,000 gallons would be carried. But it is quite improbable that more than one of these warehouses, unless located in mutual proximity, would catch fire at the same time; therefore if water mains connected one common tank with all the warehouses, a reserve of 1,000 gallons would give as great protection. This is the principle of municipal fire protection. It is the principle of centralized banking, in which the reserves of the member banks are largely concentrated in a central bank which stands ready to throw the reserve funds to the needy bank undergoing a run. It is furthermore a basic consideration in industrial combination and large scale operations, for the total quantity of reserve supplies may be much smaller when one plant feeds another plant with materials and one managing head coördinates their operations. In fact as our economic society becomes more highly organized, each part carefully articulated with the others and avenues of communication and facilities of transportation well developed, the total capital held as reserves against all sorts of business contingency may be reduced without increase of risk. The device of insurance, described later in this chapter, is closely related to these principles.

Passing risks on. In spite of all possible devices for its elimination and reduction, risk remains. Human knowledge can never attain the goal of complete prediction; the future will always be uncertain. And no human measures can wholly prevent untoward events with their consequent loss and injury. Risk is still inevitable. Under these circumstances, expedients have developed whereby those who are unable or unwilling to bear the inevitable risks of life may pass them on to others more capable and more resourceful.

Specialists in risk bearing. The entrepreneurs. The simplest example of the passing on of risks is the relation between the laboring class and their employers. As we are fully aware, production in the modern world is involved from beginning to end in risk. Many of these risks are inevitable, and someone must bear them.

But they are not, as might at first thought seem natural, shared equally by the several parties to production. The laborers especially have secured relief from many of the risks connected with production. They are, within limits, assured of a regular income, not dependent on the success of any business enterprise. This assurance is one of the characteristics of the wage contract. To be sure many risks still press upon them — the risks of unemployment, of sickness, or of industrial accident — but the risk of failure to receive their wages for work performed, because of the inability of the entrepreneur to sell the goods produced at a profitable price, is largely removed. Landowners and capitalists also, though not in the same degree as the laborers unable to bear risks, nevertheless choose not to bear any great hazards, and they also pass the risk on to the entrepreneurs.

The entrepreneur hires laborers for a definite wage, he borrows capital at a specific interest rate, he buys or rents land at a fixed price or rental. If at the end of a period his gross earnings more than cover his costs, he enjoys a residue as his own wages of management and as profit. Otherwise he fails to receive any share of the product and may even suffer loss. He thus becomes the chief risk bearer, enabling the other participants to avoid the greater part of the risks of production. For this service of risk bearing the entrepreneur looks for his reward in profits, a form of income which fluctuates widely according to the outcome of business. The entrepreneur is continually engaged in that forced speculation which, although partly transferable, is the inevitable accompaniment of production for the uncertain markets of the future. He does not buy and sell primarily for the sake of extra profits from price changes, but in seeking a normal profit on his business operations he does have to adjust his buying and selling in current markets on the basis of his estimates of the immeasurably uncertain conditions of price and volume which will prevail in future markets. His reward is subject to the risk that his estimates may prove incorrect.

The members of society automatically differentiate into groups on the basis of their willingness and ability to bear the risks of production and distribution; and one class — variously called owner-managers, business men, or entrepreneurs — offers to take

control, bear the chief risks, and appropriate what profits, if any, may result. The economy for society lies in the fact that, through this transfer to men who from the very fact of business leadership are presumably most competent to handle uncertain undertakings, risks are reduced. This competency consists in the possession of special ability, greater information and experience, sound judgment, self-confidence, and the willingness to undertake the risks of business and abide by the consequences. It is merely another instance of the division of labor. In the hands of this specialized class of risk bearers, productive enterprise is likely to be conducted with less risk of loss than if the hazards were more generally diffused.

The entrepreneurs are themselves subdivided into classes for accomplishing a further specializing of the risk-bearing function. By means of sub-contracts, risks are subdivided among those most competent to bear them in the several special fields. For instance a builder, having secured the contract to build a house for a specified sum and wishing to protect himself against contingencies about which he may know relatively little in the structural steel, stone, lumber, cement, brass, plumbing, and other trades, immediately contracts with members of these trades to supply these materials at definite prices, thus shifting the burden of risks in these lines to the sub-contractors. By this process not only the risk of the transferrer, but also the total risk of society, is reduced, because the sub-contractor is presumably wiser in his specialized line than is the general contractor.

Insurance. The most important means of transferring risk is insurance. Insurance is based on the law of large numbers. Let us assume that in a certain community there are 1,000 house owners, the houses being worth on the average $10,000 each. Each owner bears the risk of having his house destroyed by fire at a loss (on the average) of $10,000. Let us suppose further that experience shows that in this community there is an average loss by fire of one house a year. Then if each owner would contribute ten dollars each year to a common fund (plus his share of the small costs of administering the fund), it would be possible out of this fund to reimburse anyone whose house was burned. In return for a small annual payment the householder is relieved of the risk of individual finan-

cial loss by fire. Risk is passed on by each one of the group to the group as a whole.

This is an example of what is called *mutual insurance*, in which the insurers and the insured are the same. If losses are sustained which are greater than the fund accumulated against them, each one of the group is liable for his proportionate share. Insurance may also be conducted by a corporation which assumes risks as a matter of business in return for the payment of annual premiums, and whose shareholders expect the business to show a profit upon the capital which they have invested in it.

Insurance reduces risk. Insurance can be applied to any kind of hazard, provided a large enough number of independent risks can be combined, and provided that the probability of the event insured against can be actuarily determined. When these conditions obtain, it is possible for the insurance company to predict very accurately the number and amount of losses which it will have to bear in a year, and it can calculate the rate of premium which it will have to charge in order to meet these losses, pay expenses, and obtain a profit. Insurance thus applies the principle of combining risks and so reducing risk. In fact the total losses may be so regular and so predictable that the insurance company bears very little risk. While it may be highly uncertain what individuals aged nineteen will die in the United States in any year, the record of past experience makes it possible to predict with an approach to certainty how many individuals of that age will die during the year. The insurance company in writing life insurance policies for one hundred thousand people is not assuming the sum of the individual risks. In fact its risk from deaths is negligible, for it knows how many deaths to [illegible]t each year. The chief risk which it runs is that of failin[illegible] secure enough individual risks to make it possible to apply th[illegible]w of large numbers.

Note that it is n[illegible] the primary function of insurance to prevent loss. There may be as many houses burned and as many deaths as if there were no insurance. Insurance does not eliminate loss; it reduces or eliminates risk. It distributes over a group the cost of losses to that group, which cost might otherwise fall disproportionately and disastrously upon particular members of the group;

it reduces individual risk by making certain to each individual his proportionate cost on account of the statistically probable losses of the group.

When the two conditions laid down at the beginning of this section are not met, insurance in name may be nothing more than gambling in fact. There are insurance companies or underwriters that will insure against practically any future event, as the death or illness of some particular person, rain on a particular day, etc. In such cases, risk is shifted to others, but without any real reduction in amount. Insurance against drought or hail or bad weather and other similar forms of insurance are as yet hardly out of this speculative stage.

To the individual insurance appeals as a device whereby he can substitute small and certain payments (called premiums) for an uncertain and perhaps disastrous large loss. Without insurance many a manufacturer would be forced into bankruptcy if a fire were to sweep away his buildings, machinery, and materials, whereas there are but few who cannot afford the premiums in return for which an insurance company will assume the risk. The manufacturer therefore pays the premiums, and as he expects to recover them in the price of his goods he pays them cheerfully. The individual likewise regards the payments for life insurance, automobile insurance, and the like as among the costs of living, greatly to be preferred to the assumption of the risk of leaving his family unprovided for at his death or of facing a ruinous damage suit as the result of an automobile collision.

Insurance may reduce loss. Insurance is therefore a stabilizing factor in private and business life, a convenience to individuals, and a reducer of risk. Moreover insurance may even exert an influence toward reducing losses. The insurance company has of course a direct interest in reducing the losses against which it insures. And the insurers, being by training and experience specialists in their particular lines, are competent to devise measures best adapted to the avoidance of loss. Thus a company which guarantees titles to real estate is in a position to have more accurate knowledge of the situation than the purchaser of a plot of land, and it may be able to prevent losses from bad or uncertain titles.

Fire, marine, and casualty insurance companies impose elaborate rules and maintain inspection services in order to prevent fires and accidents. Even the life insurance companies have, in certain instances, inaugurated far-reaching health campaigns among their policy holders, with appreciable results in lengthening the span of human life.

Speculation in commodities. The fact that prices of particular goods and of goods in general tend continually to rise or fall has already been called to the reader's attention. To the modern business man the risk of price change, with all its possibilities of loss, is one of the most serious of the hazards of enterprise. This is a hazard of which most entrepreneurs would gladly be relieved, and there has developed in consequence a device for passing on the risks of certain price changes. This is voluntary *speculation; i.e., the buying and selling of goods for the sake of the profit from price changes.* In many lines of business there are professional speculators, men who have sufficient confidence in their ability to foresee price changes to be willing to assume the risks involved and to relieve those who prefer safety to risky profits. To be a successful speculator it is necessary to forecast the movement of prices, not absolutely, but simply better than the others in the market. Despite the fact that the existing price indicates the composite judgment of all buyers and sellers in the market, the speculator pits his opinion, manipulative skill, and luck against this judgment, and if he forecasts correctly he wins a larger return on his capital than presumably he could get for it in a line where his opinion is on a par with that of others. Faulty forecast means loss, which the speculator must bear.

The technique of speculation. The nature of the speculator's operations depends upon his judgment of the future price movement of the thing — commodity, land, government bond, corporation security, etc. — in which he is dealing. If the wheat speculator believes that the price of wheat will rise, he buys wheat now in order to hold it and sell later at the higher price; he is a *bull*, as the market jargon has it. If he foresees a fall of price, he sells now, expecting to buy later at the lower price; he is a *bear*.

It is not necessary that the wheat be immediately accepted and

held by the speculative buyer or that the speculative seller be actually in possession of wheat and make immediate delivery. Much speculative dealing is not for immediate physical transfer of the good, but for future delivery, called "dealing in futures." The bull on the wheat exchange, expecting a rise in price, buys wheat today to be delivered to him at a specified future date. When that date arrives — or before, if he wishes — he sells this wheat for delivery on the day when it will be delivered to him. He is said to be *buying long.* On the other hand, the bear, since he believes the price will fall, *sells short;* that is, agrees today to deliver wheat at a specified future date. He may not — probably does not — own today any wheat. On the specified date he will buy the wheat to deliver on his contract, or he may at any time before buy wheat to be delivered to him on the date when his selling contract matures.

Speculation equalizes prices. In any of the great speculative markets, such as the wheat market, the cotton market, the stock exchange, etc., the speculators account for a considerable part of the demand and supply of the good in question. They have therefore a profound or even predominant influence in the determination of the price. And their influence is generally in the direction of stability of price, ironing out the extremes of price fluctuation. When the price is low, presumably because the demand is slight and the supply large, the effect of the bull's buying is to increase the demand and therefore to raise the present price or keep it from falling lower. At the later date, if he has judged correctly, the bull offers wheat for sale when the price is presumably higher, and the effect of this additional supply is to cause the price to fall or to keep it from rising higher. Raising the present low price and causing the future high price to fall therefore bring present and future prices more nearly into equilibrium. The bear's operation has a similar effect on prices. He sells today, when the price is high, thus adding to the supply and tending to lower the price. At the future date, when the price is lower, his buying tends to raise it. The whole class of speculators, thus inclining to buy when prices are low and sell when they are high, have a stabilizing effect upon price fluctuations.

This relation of cause and effect is simple enough when the dealing is for immediate delivery. In the case of dealings in futures the equalizing influence, though still present, is not quite so obvious. We must remember that at the present time there are two prices, the price for immediate delivery, for example *cash wheat* as it is sometimes called, and the present price for wheat available, let us say, three months from today, or *future wheat.* Each price is the result of its own separate demand and supply, yet the demand and supply for present and for future wheat are so closely related that there is between the prices a normal *spread,* which will be maintained over considerable periods of time. When the bear for example sells wheat for future delivery he enters upon the present market for futures on the supply side, and his action has the effect of depressing the price of future wheat. The spread therefore between the prices of present and future wheat becomes abnormal, and those interested in cash wheat are thus warned that the speculators believe the price of wheat in the future will drop. Millers will feel that it is not so necessary to buy wheat now in order to be able to fill future contracts for flour, and they will diminish somewhat their present rate of buying; on the other hand, those in possession of stocks of wheat will become somewhat uneasy lest they be left with unsold wheat on their hands, and they will be inclined to be less uncompromising in their attitude toward prices. The resultant decrease in demand and increase in supply of present wheat will cause the price to drop, and present and future prices will again be in alignment. When this speculator is forced to fulfill his contracts he must enter the market to purchase cash wheat, and obviously he adds to the existing demand, thus tending to raise prices. The bear operation therefore tends to equalize prices by lowering the present high prices, and later (when the future has become the present) by raising the lower prices. Corresponding results follow the operations of the bull, who today buys long for future delivery and later sells the wheat which he has bought.

Professional speculators are more apt to be right than wrong in their speculative judgment, but they are not infallible. If their speculative judgment is generally wrong at a particular time, price variations are of course exaggerated. At times when the amateur

public has been attracted into speculative operations, unrestricted speculative activity, even though based on sound professional judgment, may, because of the cumulative influences of mass psychology, tend temporarily to cause excessive variation of prices in a stampeding market.

Speculation promotes economical consumption. As a corollary to the equalization of prices, speculation performs another important service. When goods are abundant and therefore cheap, the speculator's influence tends to take goods from the present market and hold them for sale at a time when relative scarcity is anticipated. Present consumption will be somewhat curtailed in order that there may be more ample supplies for the future time. Without the operations of the speculator plenty and famine are likely to follow each other, but with more or less accurate forecasting of price changes the risk of famine is almost entirely avoided. In other words the speculator creates time utilities when he carries goods over to a period when the demand is more urgent. One of the earliest examples of speculative wheat dealing is seen in the dealings of Joseph in Egypt. During the years of plenty he accumulated stores of grain and other provisions to tide the Egyptians over the seven lean years, doubtless with personal profits commensurate with his shrewdness in applying his "inside information" as to price changes. In the same way speculation makes it possible to utilize present stocks more fully and thus avoid carrying goods over from a period of scarcity to a period of relative abundance.

Hedging. In addition to its simple use as thus far considered, speculation offers a peculiar method of risk avoidance, known as *hedging*. Two illustrations of its use — by a miller and by a grain dealer — will serve to disclose its essential features.

Suppose a miller of Peoria, having noted the cost of materials, labor, and equipment for milling, contracts on June first to deliver to a jobber on July first 1,000 barrels of flour at ten dollars a barrel. The ten dollars are sufficient to cover the cost of milling and delivering the flour together with a fair profit, provided the price of wheat does not rise. Of course the miller might buy at once and hold the entire quantity of wheat required for making this flour.

But the miller is seldom in a position to do this; he may not have adequate storage facilities, and he may not find in the samples on June first the exact kind of wheat he needs. The former obstacle might be surmounted by taking advantage of simple speculation and buying wheat for delivery at future dates as required by his milling operation, but the difficulty of finding exactly the right samples would remain. For these and other reasons, millers generally prefer to buy their wheat in moderate quantities from time to time as they need it. But this involves the risk of loss from a rise in the price of wheat. In this situation, the miller has resort to *hedging, the essential feature of which is the elimination of risk by taking an equal position on both sides of the market.*

Our miller, on the very day he contracts to deliver the flour, buys on the Chicago Board of Trade 5,000 bushels (by assumption enough to make the 1,000 barrels of flour) of July wheat. Of course he does not intend to grind the July wheat; it is not deliverable till during July, whereas the flour is due July first. He buys it because it affords him a hedge until he does find occasion to buy his milling wheat. He waits, let us say, until June sixteenth, when on the basis of samples he buys 5,000 bushels of milling wheat and instantly sells 5,000 bushels of July wheat. Now if, as is the usual case, cash wheat and July wheat differ by a normal spread, say four cents, when one rises or falls the other does likewise, staying four cents apart. Therefore if by chance the price of cash wheat rose 13 cents between June first and sixteenth, the miller's costs would increase by 13 cents a bushel above the figure he used in calculating the price at which he agreed to deliver the flour. But it is also plain that the July wheat also rose 13 cents above its cost by the time it was sold June sixteenth. Therefore the excess which he has to pay for his milling wheat he has gained on his July wheat. If the price had fallen he could have bought his milling wheat lower but would have lost on his July wheat a similar amount. In this manner, whether the market price rises or falls, his milling profit is protected.

Turning now to the second case, we may assume that an elevator at Wichita ships grain to Galveston directly, a distance requiring 20 days' time, but hedges on the Kansas City exchange. Let us

suppose that the normal difference between Wichita prices and Galveston prices is 15 cents — a difference sufficient to cover the cost of shipment from Wichita to Galveston plus the grain dealer's profit — and that prices at Kansas City are likewise higher, say by 8 cents, than the Wichita prices and for the same reason. Let us suppose also that on January first May wheat was higher than cash wheat by 5 cents, a difference caused by the cost of carrying. These being the conditions on January first, the Wichita elevator notes that on this day the price of cash wheat is $1.00 in Wichita and $1.15 in Galveston. With this difference, there is a fair profit in shipping wheat to Galveston. But, since the wheat cannot be sold there until it arrives, about the twentieth of January, there is the risk of a fall in price which might wipe out the profit or turn it into a loss.

To eliminate this hazard, the elevator hedges. It ships, let us say, 10,000 bushels of wheat to Galveston and immediately sells on the Kansas City market 10,000 bushels for May delivery at a price of $1.13. The grain arrives at Galveston on January 20 and, wheat prices having in the meantime risen by eighteen cents, is sold for $1.33 a bushel. Immediately the Wichita elevator buys on the Kansas City market 10,000 bushels of May wheat to cover its short contract, paying $1.31 per bushel.

The price situation in this example may be better visualized with the aid of the following table.

	Wichita		*Galveston*		*Kansas City*	
	Jan 1	Jan 20	Jan 1	Jan. 20	Jan 1	Jan. 20
Cash wheat . .	1.00	1 18	1.15	1.33	1.08	1.26
May wheat . .					1 13	1.31

The results of the hedge are as follows:

(*a*) On the May wheat:		
Selling price, January 1		$1.13
Cost, January 20		1.31
Loss		.18
(*b*) On the wheat shipped:		
Possible selling price at Wichita,	Jan. 1	$1.00
Actual selling price at Galveston,	Jan. 20	1.33
Gain		.33

Net gain from both transactions, 33¢ − 18¢ = 15¢, which by assumption is the gross margin at which the Wichita elevator can handle Galveston shipments and make its regular trade profit, and which was the margin on which the elevator made its decision to ship on January first. If the price of grain had fallen during the time of transit, the elevator would have lost on the grain shipped but have gained on the May future.

The essential features of hedging are thus brought to light, especially these three: (1) no profits are made or are meant to be made from the hedge itself, (2) it is employed to protect a trade profit — the miller's for milling flour, the grain merchant's for dealing in grain — against loss occasioned by change of price, and (3) it depends upon the existence of a normal spread between the prices prevailing in two markets, the prices of two grades of the same commodity, the prices of convertible stocks and bonds, etc.; if this normal spread is not maintained the hedger may lose or gain by the amount the spread departs from normal.

Social service of speculation. Manufacturers, merchants, and entrepreneurs in general find in organized, voluntary speculation a welcome relief from the hazards of price changes. The development of professional speculators is another example of the division of labor, which tends to increase the efficiency of our society Relieved of some of the risks of business, the manufacturer or merchant can devote more time to the acquisition and employment of skill in his own specialty.

Moreover the transfer of risks from the relatively ignorant and uninformed to men who are experts in their particular fields means a reduction in the hazard. Less risk attaches to a contract for the delivery of wheat at a given price at some future date made by an able speculator than to a similar contract made by a small local grain dealer, for it may be presumed that the speculator has the wider and more profound knowledge of world conditions affecting the demand and supply of wheat

More than this, speculation, by equalizing prices, tends to do what insurance also does, namely, diminish the very events against which it operates. And finally there is the economy in consumption over long periods of time which speculation promotes. All

these are services to society of great importance, well worth the price paid the speculating class in the form of speculative profits.

Investment, speculation, and gambling. That voluntary speculation is not an unmixed blessing to society is principally due to the fact that the machinery of speculation is available for the gambler as well as for the speculator.

It is not always easy to distinguish between the three concepts of investment, speculation, and gambling. The difference between speculation and investment lies very largely in the attitude of mind toward risk. The speculator is willing to assume the risks involved, whereas the investor prefers a moderate profit with the least possible risk. Inevitably he must assume some risk in any investment, but he reduces this to the minimum. The investor therefore tends to buy the securities of a stable government or the first mortgage bonds of well-established corporations or stocks of established reputation, and he looks for his gain primarily to the annual income of interest or dividends The speculator, on the contrary, is attracted by the stocks that are likely to fluctuate in price and, thinking less of the income from dividends, buys or sells as he thinks the price is going up or down.

As we have seen, successful speculation is of advantage to society as a whole through the reduction in the volume of risks which it effects. Gambling, on the contrary, is marked by the creation of artificial and unnecessary risks. Since one gambler gains what another loses, gamblers as a class cannot gain. Within the class, the especially shrewd and the dishonest gain at the expense of the others. The average outsider, betting against experts and professionals, has little chance of success at gambling, whether on horse racing or on price movements in the pit, the stock exchange, or the realty market. The social verdict which considers the gambler a parasite on society is probably fully justified.

The corner. It has occasionally happened that a few powerful speculators have secured sufficient control of the supply of some commodity to enable them to determine the price. This was one of the real evils of speculation, but it is one whose importance was greatly overestimated by the public. Successful corners have been very rare and have lasted only a few days at the most. Price control

through corners has had little effect on the public, as the corner was usually aimed at other speculators who had made contracts for future deliveries and who were forced to buy the commodity, whatever the price, in order to fulfill their contracts. The occasional spectacular winnings of those who have engineered successful corners were obtained from the other speculators and were not a cause of loss to the general public, except as its members voluntarily took gambling chances on the market. Corners are now prohibited in the United States by exchange regulations.

Judgment of speculation. Some of the most favored and persistent of the economic fallacies which popular opinion cherishes centre about voluntary speculation and the speculative market. These fallacies generally rest upon ignorance of the laws of price, though they are to some extent excusable on account of the difficulty, already noted, of distinguishing always between investment, speculation, and gambling. When it is charged that a great part of the dealing on the speculative exchanges is gambling, the truth of the charge must be conceded. However the seriousness of this evil, from the public point of view, is generally much exaggerated.

The spectacular profits sometimes acquired by successful speculators have had a dazzling effect and have fostered the belief that the speculators as a class are amassing huge fortunes out of the producers on the one side and the consumers on the other. As a matter of fact, extraordinary speculative profits are generally at the expense of equal losses by other speculators, and it is not at all probable that the net income of the whole class of speculators is more than a fair reward for the useful services which we have seen that they perform.

The most persistent charge is that on the exchanges shrewd and unscrupulous speculators are able to fix prices at their own sweet will, "in violation of the law of demand and supply" and to the enormous injury of both producers and consumers. Waiving the obvious difficulty of understanding how the same price fixing can injure both producers and consumers, the fact is that nowhere in the world is there so close approach to determination of price by demand and supply according to economic law as in the great spec-

ulative exchanges. It is here that we must always turn for our practical illustrations of the theory of price under conditions at least approaching theoretical pure competition. Except in the rare and brief case of a corner, it is quite impossible for the speculators to "rig" the price, for the simple reason that two groups, with opposed interests, are always working against each other. The speculative market offers the minimum of friction to the operation of the laws of demand and supply.

The opponents of speculation often go to the length of demanding extreme restrictive legislation or regulation, the prohibition of "futures," or even the entire abolition of the speculative exchanges. There are aspects of speculation which justify public regulation, but any action which would seriously interfere with the normal operation of voluntary speculation would be a costly blunder. An impartial weighing of its merits and demerits can leave no question as to the net advantage to society as a whole of voluntary speculation in commodities. The unavoidable speculation which the prevailing economic system forces upon nearly all business men stands as an unfortunate necessity.

An interesting illustration of these truths occurred at the time of the Civil War. "When the greenbacks became the regular money of the country, there was established in the city of New York a 'gold exchange,' a speculative market where people met daily to buy and sell gold, and where the price of gold for the whole country was determined. As the greenbacks depreciated, there were great complaints about the high price of gold, and many people charged that the speculation on the gold exchange was responsible. This was exactly the same sort of complaint that we hear from time to time now about the New York Cotton Exchange or the wheat speculation on the Chicago Board of Trade. Congress gave ear to these complaints and passed a law abolishing the gold exchange and forbidding all speculative dealings in gold. The result was immediate and disastrous. Nobody now knew where he could get gold or sell gold or what the price was. There was now no common point where all the supply and all the demand could come together and determine the price. Instead there was bargaining and higgling at various places, and widely differing

prices were quoted in different places. The price, instead of falling as hoped, immediately rose higher than ever. It was 198 on the day the law was passed. The next day it was 208; the next 230; and in less than two weeks it had risen to 250. This was the natural result of the bidding of people who had to have gold and who no longer knew where to get it or what was a fair price. The demand for the restoration of the exchange was so prompt and so strong that Congress repealed the law just two weeks after its passage. From that time on the gold exchange did business till the greenbacks came to par in December, 1878." [1]

[1] Fairchild, F. R., *Economics*, 1936, p. 370.

XXXIX

THE BUSINESS CYCLE

Business fluctuations: In general. In a developing social order economic forces do not remain for long in a state of stable equilibrium. Changes are constantly occurring in some branches of activity which do not instantly harmonize with corresponding changes in other branches. Forces determining the demand for a given good operate independently of the forces determining the supply, with the result that the market for this good is disrupted until a change in price can bring it again into balance. The tendency for capital and labor to be distributed among the different productive enterprises so that their marginal productivities are everywhere the same is thwarted by unforeseen changes of fashion, by invention, by political disturbances, and similar events. The price level may be shifted by the forces of change, and the individual prices within the system which were formerly in approximate balance with each other may move erratically, sometimes even in opposite directions. These are random examples of economic alterations which tend to keep our productive system in a state of disequilibrium and uncertainty.

It is impossible of course to analyze each of the countless changes in economic activity which characterize our social order, but there are some major aspects of this problem which do repay study. When we examine the course of economic activity over a period of years we find that the changes are not altogether sporadic and erratic. On the contrary, there are certain well-marked fluctuations affecting the productive system in its entirety. Business in general is more active at certain times than at others. The total volume of production is seen to increase during one period of years, whether this be measured in absolute quantities or in value terms, and to fall during a subsequent period. Sometimes there are jobs aplenty; later there is widespread unemployment. Boom times alternate with times of business depression. These general fluctuations of

economic activity must be the result of forces which play upon the productive system in a comprehensive manner. It is our purpose in the present chapter to analyze these fluctuations and so far as possible to ascertain their causes.

Types of economic fluctuations. When we speak of an economic fluctuation we mean a change of business activity on a fairly broad scale which can be represented in statistical terms. The statistical index chosen to represent the change may be selected from among a number available. For example we may chart the volume of agricultural and factory production over a period of time, or the amount of employment of wage earners, or the average price level, or the rate of return on invested funds, and find definite evidence of a general movement upwards or downwards as the case may be. In the following pages we shall use such indexes as these as illustrations to aid us in the analysis of the problem of economic fluctuations.

Economic fluctuations may be classified in three main categories: (1) independent irregularities; (2) cycles; (3) the secular trend. Of these types of fluctuation, the second or cyclical type is by far the most important, both as a matter of practical concern to mankind and as a problem of economic theory. In this chapter therefore we shall give our attention almost exclusively to the cycle, pausing only to note the existence of the other two forms of fluctuation.

Independent irregularities. This term is used to designate such disturbances in the normal development of economic enterprise as result from isolated events, such as widespread strikes, national disasters, or popular reactions to changes in governmental policy. These disturbances are usually short-lived and self-correcting, but their effects appear on any long-run chart of business activity as irregularities not to be accounted for by reference to the major forces of change which are at work.

Cyclical fluctuations. As the term indicates, a cyclical fluctuation is one which discloses a recurring, rhythmical character. During a complete cycle business activity will pass through a phase of acceleration followed by a second phase of decline. In a series of cycles this process will repeat itself over and over again, though not

of course in exactly the same statistical terms. Long-range studies of business trends have brought to light not one but many different cyclical movements which seem to be characteristic of our economic order. These cycles may be classified in four important groups: (1) seasonal cycles, (2) long-wave movements, (3) minor non-seasonal cycles, (4) major non-seasonal cycles.

Seasonal cycles. These are the simplest and most readily identified of the business cycles. They result in part from market response to the changing climatic seasons, in part from the conventions of our calendar. The increased business activity which corresponds with seedtime and harvest in an agrarian country is an example. So also is the business rush during the Christmas season as contrasted with the slack markets of midsummer.

Figure 3 on page 219 has been prepared primarily to illustrate other types of cycles. However, reference to it will provide the student with evidence of the seasonal fluctuation of one index to business activity: *i.e.*, the trend of factory employment and payrolls. Disregarding changes which occur over a period of several years, it will be seen that each year considered by itself records a definite seasonal variation. The line indicating payrolls, for example, tends to reach a high point at two different times within the year and again to fall to a low point at two other times. This is more clearly shown in some years than in others — in the year 1926 for example. Still the tendency can be detected in every twelve-month period covered by the chart.

Seasonal fluctuations do not present a major social problem. It is true that certain specific industries — such as the manufacture of clothing, coal mining, the building trades — may be seriously affected by them. In such industries the problem is one of stabilization through improved business organization and management. In society generally the existence of seasonal fluctuations is clearly recognized, and economic activity as a whole is fairly well adjusted to them.

Long-wave movements. There is general recognition of an important group of cycles, somewhere between forty and sixty years in duration, which are usually designated as long waves. Their presence has been detected in the trend of price data and production

data for the United States, England, and France. Figure 2 reveals the character of the long waves in wholesale prices in the United States, with low points in 1843 and 1896 and high points in 1814, 1864, and 1920. It is important to note that this chart is drawn to a *ratio scale* and hence shows proportional changes rather than absolute rates of change.[1] An analysis of long-wave movements falls outside the scope of the present discussion, but we may note that the underlying causes of these long waves have much in com-

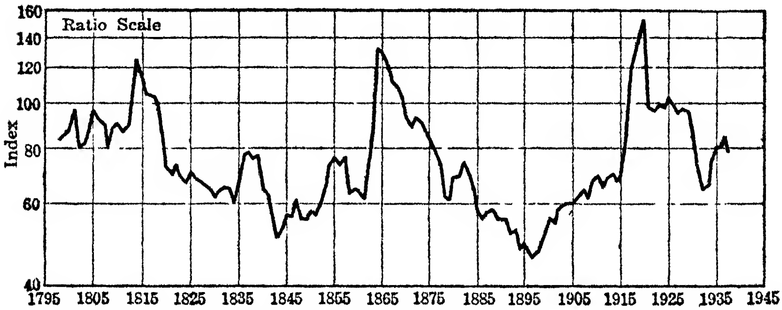

FIG. 2. Index of Wholesale Prices, All Commodities, United States, 1795-1938.[2] (1926=100)

mon with those of business cycles proper and that much of the ensuing discussion of the causes of the business cycle may be extended to explain this slower fluctuation in business activity.

Major and minor cycles. The term *business cycle* is ordinarily used to refer to fluctuations in employment and output which occur with an average duration of approximately twelve years, but this description is not quite adequate because the experience of the past twenty or thirty years reveals that there are in fact two business cycles — a major cycle and a minor cycle. Reference to Figure 3 will show that factory employment and factory payrolls rose from

[1] The ratio scale is drawn so that a given vertical distance always represents the same proportional change, rather than the same absolute amount of change (160 is the same distance above 80 as 80 is above 40 and as 120 is above 60, etc.). Lines drawn against a ratio scale show directly rates of change; a line continuing straight showing a constant rate of change; a line concave to the horizontal axis shows a declining rate of increase or an acceleration of rate of decrease, depending on the direction of movement, while a line convex to the horizontal axis shows the reverse; a line rising or falling more steeply than another shows a faster rate of increase or decrease, respectively.

[2] Warren and Pearson index to 1889; United States Bureau of Labor Statistics from 1890.

a low point in 1919 to a high point in 1920 and then dropped to a new low point at the end of 1921. Beginning here the process repeats itself, another high point being reached in 1923 and another low point in 1924. Then a slow rise to a peak in 1929 and an abrupt fall to a very low point in 1933. This is a record of a minor cycle, in this case varying in duration from 2½ to 5 years. Another glance at the chart will expose a longer rhythmical swing, particularly with regard to the points of deepest depression. An extremely low point

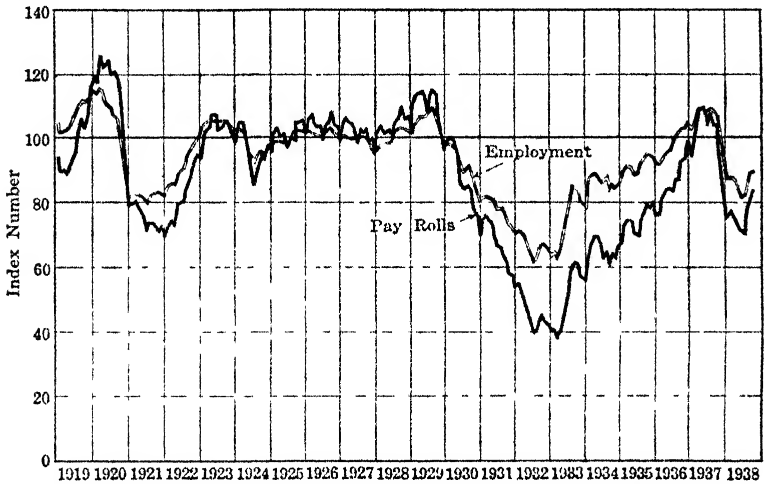

Fig. 3. Employment and Payrolls, All Manufacturing Industries.[1]
(1923 25 = 100)

was reached in 1921, and another in 1933; this is a major cycle embracing a twelve-year period. The high points of 1920 and 1929 illustrate the opposite phase of the major cycle, in this case a period of nine years. The chart we are using covers too brief a period to show these major cycles to advantage; but practically all long-term indexes of economic activity reveal major cycles, recurring in periods of from 7 to 12 years, and running concomitantly with briefer minor cycles.

It is impossible to say which of these two types of cycle will be the more important in the future, but in all probability the longer business cycle of seven to twelve years' duration is the one that

[1] United States Bureau of Labor Statistics, *Employment and Payrolls*, October, 1938, p. 19.

deserves closer study. The reader should realize however that business cycle analysis and any discussion of the causes of these fluctuations in employment may be applied to the explanation of either the shorter or the longer variation.

The secular trend. One other type of variation must be mentioned before we discuss in detail the business cycle proper. Even in the

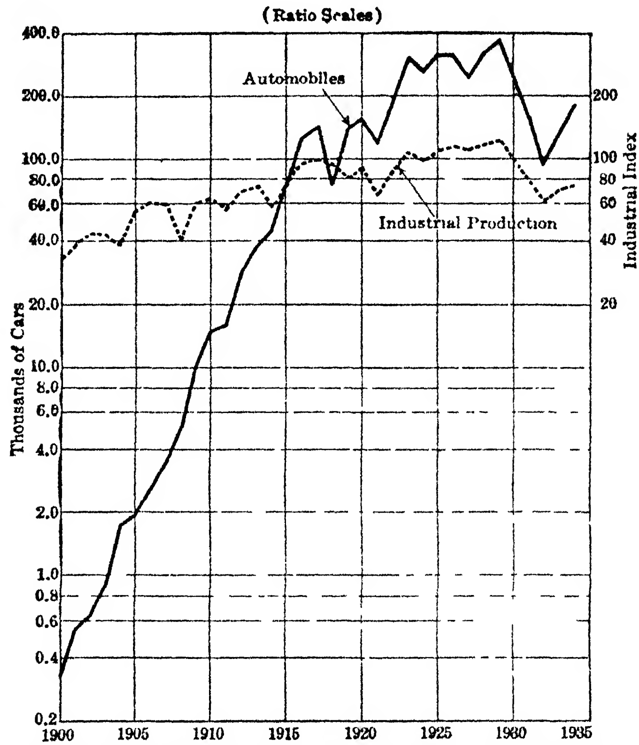

FIG. 4. Average Monthly Production of Passenger Automobiles and General Index of Industrial Production, United States, 1900–1934

absence of cyclical fluctuations our productive system would not remain entirely static. There are forces of growth which would carry economic development in a given direction, even though the different branches remained in equilibrium with each other during the process. The term secular trend refers to this normal line of balanced movement. In the broadest meaning of the term this trend of development may be either upwards or on a uniform level

or downwards, and specific industries can be found which illustrate each of these trends during any period of time of reasonable length.

If students of cyclical change do not take account of secular trend, their statistical image of the cycle may be badly distorted. This may be illustrated by reference to Figure 4. It will be seen that the trend of automobile production during the period 1900–1919 was almost steadily upward, although the chart of industrial production as a whole shows clearly defined cyclical movements during this period. Obviously a statistical study of any group of industries in which automobiles formed an important part would be strongly affected during these years by the secular trend of this specific industry. The important point for our purpose is to note that if the figures are to be a faithful measure of cyclical change alone, they must be corrected for secular trend; otherwise they will reflect not only the effects of the cycle but other forces of growth as well. It is our intention merely to make this clear, not to discuss the methods of measuring secular trend and eliminating its influence on statistical series. In the discussion which follows and in Figure 5, which illustrates the business cycle, it may be assumed that the necessary correction for secular trend has been made.

The business cycle: General description. Each of the various types of economic fluctuation which we have mentioned exerts its own effect at any given time upon business activity. Furthermore none of these variations affects all departments of business activity in precisely the same manner. Different strands of business activity may vary during a given cycle by as much as a year in the times at which they reach the peak and the trough of a given depression. When dealing with these cycles therefore we have to do with very complex phenomena, a large number of movements in economic life, which are partly independent of each other and partly interacting.

Despite this complexity, it has become customary to speak of "the business cycle" as if a single general movement took place over a term of years affecting in the same direction the whole range of economic activity. Such movements, if they exist, must result from the interaction of all the major and minor cyclical tendencies operating in the period. That the "business cycle" in this compre-

hensive sense is an actual phenomenon will become clear upon examination of Figure 5. This is an index of general business activity in this country over the past sixty years. It will be seen that, apart from interim disturbances, the trend of business activity as a whole, resulting from complex and diverse cyclical movements, describes a definite pattern within certain well-marked periods of fluctuation: 1878–1885; 1885–1897; 1897–1908; 1908–1915; 1915–1921; 1921–1933. Although the existence of different cyclical movements within the business cycles must be kept in mind, there is practical justification for taking a simplified view of the problem of fluctuations such as is implied in this generalized concept of the cycle. For these are the variations of business activity which are

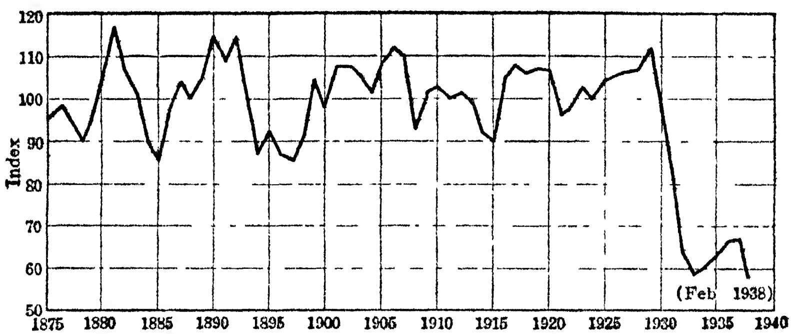

FIG. 5. Index of General Business Activity, United States, 1875–1938 [1]

really felt. They contain the major features of economic disturbance which present business men as a group with weighty problems and throw on society the strain of widespread maladjustment.

Phases of the business cycle. When we compare the series of cycles pictured in Figure 5, we find that they were in many respects dissimilar. For one thing, they were not of the same duration, as one can see by comparing the period 1878–1885 with the period 1885–1897 and so on throughout the chart. In other words, no standard period of years can be designated as the normal duration of a business cycle. Nor did these cycles correspond in degree of fluctuation; there was greater fluctuation in the 1885–1897 period, for example, than in the 1908–1915 period. But despite these

[1] Clearings index of the Federal Reserve Bank of New York, with adjustment for secular trend, seasonal variation, and changes in the general level of prices.

diversities there still remains a recognizable pattern of similarity. They all passed through four rather well-marked phases, to which it is customary to give the names, revival or recovery, prosperity, recession, depression. The revival phase is taken as the starting point of the cycle; for example, the 1878–1885 period and all subsequent periods begin when the indexes of business activity turn upwards. The prosperity phase is marked by the later stages of this upward trend. Recession is defined by the downward slump of the curve. Depression is indicated by the subsequent low points on the curve, continuing until another revival sets in. With minor differences in detail this general pattern is seen to apply to all business cycles. It is now generally believed that all these successive phases of the cycle, except the first, are self-generating. The revival phase develops through its own momentum into the prosperity phase; the prosperity phase contains within itself forces which initiate recession; recession in turn develops through its own characteristics into depression. The phase of revival is an exception to this rule. Depressions are not self-limiting; they seem incapable of moving into a new revival without the impact of some external force.

The conditions prevailing in the successive phases of a cycle may be roughly summarized. In revival we find rising prices of securities and commodities, a rising volume and value of industrial production, increasing sales and forward orders in most lines of trade, increasing employment, mounting loans or debts, and a rising rate of interest. In prosperity we find prices continuing to rise but gradually leveling off, large accumulated stocks of goods, high but stabilized employment, an intensified demand for credit accompanied by high rates of interest, and the production of capital goods beginning to weaken. In recession the business picture is one of decline in sales, orders, production, employment, and prices, an increasing real burden of debt as the amount of debt is reduced more slowly than the decline in the value of assets, and a falling rate of interest. In the depression phase debts and stocks of goods have been reduced; prices run at low levels, though security prices may show some slight return of strength; the rate of production is low; unemployment is widespread; interest rates are low. The method

of development of these conditions is the important problem for our consideration.

Revival. The most significant feature of the revival, which we take as the starting point of the business cycle, is an increase in spending. Ordinarily we find that business spending increases first and is followed subsequently by an enlargement of consumer spending. Economic science has as yet no complete answer to the question as to what the forces are which induce this increase in spending at a time when business is depressed and unemployment is widespread. There are however some points on which there is general agreement.

In the first place, depression itself to some extent serves to prepare the way for a revival. Weak firms and weak banks which have survived the first months of the depression will eventually be forced through bankruptcy and dissolution or reorganization, and the surviving firms, with assets written down and capital structures reorganized, will be better credit risks and so in better position to secure funds for spending whenever business spending appears to be warranted.

Furthermore the continued deflationary pressure of the depression itself is almost sure to bring about a reduction in costs. During a depression wage rates, interest rates, and many of the prices of producers' goods will have been reduced. Not only are costs, both of current production and of new investment, lowered in this fashion, but the less efficient members of any working force and the less efficient producers in any given industry are likely to be the ones first thrown out of employment. The favorable effect of these reductions in cost should not of course be exaggerated. For one thing, it is quite clear that most reductions in costs bring about equivalent reductions in the incomes of other members of the economic system, so that unfavorable effects on purchasing power may tend to counteract the favorable effects of cost reductions. Nevertheless if the depression has the effect of decreasing certain sorts of costs that have been out of line with the general level of prices and incomes of the community, it may make the prospects of future profits more favorable and thus encourage increased business spending.

But more important is the fact that, if business is to maintain even a low rate of production, the time will come when it will require new machinery and equipment to replace that which has worn out. During a period of depression it is the practice of business firms to reduce expenditures for maintenance of plant to the minimum by postponing replacements of worn-out equipment, but this process of using up capital goods cannot continue indefinitely; at some time increased spending for this purpose will be inevitable.

All three of the causes of revival so far discussed are in some degree automatic in their operation, and they have considerable significance. But any history of the business cycle makes it clear that the most powerful forces working to initiate revival are ones which do not develop automatically out of the preceding phase of the business cycle but are in a sense fortuitous. Of these the most important is the appearance of new investment opportunities.

Historically the classic example during the nineteenth century was the development of railroads throughout the world. In the United States, up to 1890 at least, this provided the largest single outlet for investment funds. The existence of such investment outlets affected not only the industries directly and indirectly concerned with the railroads themselves; the high rate of investment in railroads made for a high national income and thus stimulated investment in many other lines. A more recent example is the automobile industry in the United States, which during the decade of the 1920's grew at a prodigious rate. Though founded many years previously, the industry had evolved by 1921 to the point where it presented an ideal opportunity for the investment of capital on a large scale. This investment opportunity and the shortage of housing that had developed during the war were largely responsible for the quick ending of the 1921 depression.

Investment opportunities may take several forms. Though perhaps the most important ones occur when a new industry or a new area begins to develop, much investment is undertaken by existing firms in existing industries and without reference to the opening up of new territories anywhere in the world. Such investment frequently occurs as the result of new and cheaper methods of produc-

tion, such as are being discovered almost continuously. After a depression has been in progress for several years, even if no major new investment opportunity appears, replacement of old equipment and expansion of plant are eventually forced upon business men both by the wearing out of what they already had and by the availability of new kinds of machinery and new productive methods.

Revival breeds prosperity. Once a revival is under way it tends to become cumulative, feeding on itself for a considerable period at least. It is not difficult to see why this is so. As new construction projects are launched and carried through, men will be employed in building factories, installing machinery, and other such operations. Part of the increase in expenditure will thus be matched by increases in wage payments. Many of the new expenditures will go toward the purchase of building materials, machinery, tools, etc., and the business firms which supply capital equipment will thus find the demand for their goods increased and their incomes enlarged. They in turn will find it necessary to place orders for raw materials, to engage a larger laboring force, and to revise their production schedules upward.

Consumers who have benefited from the primary increases in income will of course increase their expenditures for consumers' goods, and the demand for these goods will be swelled still further as the larger consumer incomes generated by renewed activity in the capital goods industries manifest themselves. As wage earners come to feel that their jobs are secure and their incomes likely to increase, they incline to expand their expenditures beyond their immediate means and so resort to borrowing in one form or another. This factor is of great importance for those industries which produce durable consumers' goods, such as automobiles, refrigerators, washing machines, and the like, and especially in the field of housing. During the decade of the twenties consumer spending for the purchase of new houses was of very large dimensions and was largely financed by borrowing from the banks. This substantial increase in the demand for consumers' goods of course gives rise to increased production and employment in the consumer goods industries and still further reinforces the trend toward increased business activity.

Furthermore investment in stocks, both of raw materials and of finished goods, is certain to be stimulated. During a depression business men almost invariably attempt to reduce their inventories to the lowest possible level, but once revival is under way this entire picture changes. The expectation that the volume of sales will continue to rise and the fear that a shortage may develop and that deliveries of goods may be slow will induce business men all along the line to build up much larger inventories than they have been carrying. Rising prices of goods of all sorts give them an added incentive, for even though they may be most anxious to avoid speculation in inventories, the expectation of a continued rise in prices will make it seem prudent to add substantially to the volume of goods carried in stock. The initial increase in the demand for both producers' and consumers' goods tends thus to be magnified.

As recovery proceeds, and especially if it proceeds rapidly, the increasing demand for the use of the factors of production and for both consumers' and producers' goods will bring about higher prices, but it is characteristic of a period in which the general level of prices is rising that many of the prices which determine costs of production — such as wages of labor and prices of capital goods — lag behind the sale prices which determine business income. The result is a wider margin of profit, which serves as an added stimulus to expansion and employment and therefore to income.

In summary we may say that the initial increase in spending gives rise to a primary increase in income. But this in turn stimulates an increase in production over wide areas of the economic system, and the increase in production in turn gives rise to a secondary increase in income. Even there of course the process does not stop, for the enlarged output of capital goods gives rise to an increased demand for the raw materials which they require and hence to further increases in income. Thus the initial increase in spending spreads throughout the economic system, and as it spreads the level of production rises, consumers' and business incomes are increased, further spending follows, and production is again stimulated.

Limiting factors in prosperity. What are the limits to this process? There would of course be none if each expansion of spending and income gave rise to a further expansion in spending of exactly

the same size. For in that case the assumed initial increase in business spending would give rise to an infinite series of further expansions, each of the same size. But actually each successive expansion will generally be smaller than the one preceding it. It is to be presumed that consumers will plan to save at least some portion of the increments to their incomes, and it is highly probable that each increase in incomes will give rise to some stable increase in the rate of saving. This means that consumers' spending will normally increase less than consumers' incomes. Now this would not in itself argue a declining rate of increase in total spending if consumer savings were, as is normal, made available to business and spent for investment in new plant and the like. Similar considerations however apply to the expansion of business incomes and business spending. If a firm receives a rush of new orders, its disbursements in the form of wage payments and payments for raw materials will of course increase, but it is most unlikely to begin paying out higher dividends the moment higher profits begin to be earned. Business men, like consumers, will plan to save some part of the increment of their incomes in the form of liquid assets. This implies that, although total spending in any such period will be increasing, it will be increasing at a decreasing rate.

This fact will be reflected in the relative movement of prices and costs. After the initial upward surge, industrial prices tend to rise for a period at a decelerating rate. In the meantime some of the more sluggishly moving prices which are costs to industry, such as wages, interest, and the like, begin to catch up and to decrease the margin of profit. Increasing costs weaken the motive for further expansion of capital equipment, and the same effect comes from the tendency of increased capital resources to show a declining physical productivity at the margins of application. The expansion of the capital goods industries and of their disbursements is retarded, tending to a decrease in consumer income and a slowing down of the rate of increase in purchases by consumers and retail dealers. Prices of consumers' goods then show fluctuations around a temporarily stabilized level or even show a slight tendency to decline.

Recession and depression. At the culmination of a period of prosperity the developments we have traced finally interact so as to

bring a reduction in the rate of prospective profit. The profit margin is pinched between a retarded rise of selling prices in general and an accelerated rise of costs. As the rate of interest finally rises above the rate of prospective profit, the incentive to borrow for expansion is replaced by an incentive to curtail business spending and pay off loans. As soon as the industries producing capital goods find a definitely declining market, they add an absolute reduction of output, disbursements, and bank loans to the now rapidly changing picture of business. With consumer spending then further curtailed, a definite decline in industrial activity becomes general, and profit margins are further narrowed as business costs tend to fall less rapidly than other prices. Thus begins a spiral of falling values, forced sale of goods to permit required debt liquidation, consequent reduction of credit currency, further depression of values, increased unemployment, and a general decline in incomes and trade. A psychological depression results, and a reduced velocity in the turnover of currency augments the other factors. Declining prices of securities add further to business curtailments, partly by their psychological influence, partly because of their effect on the credit standing of business enterprises.

As trade approaches a minimum level and surplus stocks of goods are exhausted, as the burden of old debts is slowly removed, as the rate of interest falls, as costs of production fall into line with the general decline of prices, profit margins tend to increase enough to permit the use of funds more profitably than in paying off loans. Such new funds might be used for slight expansions here and there in the volume of operations or for meeting the requirements of postponed maintenance of capital equipment. A period of uneven but general equilibrium between costs and selling prices, between production and consumption, thus tends to prevail until events external to the conditions of depression occur to stimulate another revival.

Forces initiating the business cycle. Apparently the conditions internal to the business situation in a period of depression can at most set the stage for a possible revival, leaving it to external forces to set the cycle in motion. The normal operation of the credit system exerts no original impulse upon the cycle; low interest rates, excess bank reserves, and super-abundant supplies of credit proved

quite ineffective in 1932 and 1933 to bring revival out of depression. Similarly the equilibrium of costs and income which obtains throughout the business world after recession has worked out its effects is only an invitation to accelerated activity. The spur of self-interest under such conditions certainly moves enterprisers constantly to seek opportunities to expand their businesses. These impulses may result in a degree of growth on a somewhat uniform level, but they do not appear to be effective for starting the upward accelerated surge of activity which is characteristic of a cycle in its first phase.

As a rule depressions have not ended until substantial new investment outlets have been discovered. These may take the form of government borrowing for war or for a program of armament, or the growth of a new industry, or the opening up of new territory, or a general advance on a number of fronts simultaneously. The necessity of making up arrears of depreciation is often a powerful contributing cause of the up-turn and may even have a good deal to do with its exact timing, but as a rule we must look to some cause external to the operation of the economic system itself in order to explain the initiation and continuation of a prolonged upward movement.

Consequences of cyclical fluctuation. Generalizing from past experience of cycles, it may be said that business activity fluctuates as a rule between twenty-five per cent above and twenty-five per cent below its average level. This fluctuation is in addition to the changes produced by long waves from one generation to another, these long-range changes being interpreted as shifts in the average level about which the cycle oscillates. The percentage of fluctuation given does not apply of course to all industries and all occupational groups alike, for in any business cycle some interests are much more seriously affected than others.

These continued fluctuations of the economic system have many injurious social consequences. Most obvious is the general impoverishment which occurs at the depression points, a temporary phase for society as a whole, but permanently disastrous for multitudes of individuals. But aside from this the alternations between general plenty and general want are seriously disruptive to the social order.

They upset the morale of the people and develop a fear psychology by introducing uncertainty and insecurity into their lives. Experience of recurring business depression, with its reduction of wage income and widespread unemployment, has its effect on the efficiency of wage labor. These wide fluctuations upset business leadership by thwarting the best laid plans for future economic development. They draw the various branches of industry out of adjustment to each other, thus diminishing the efficiency of individual enterprises and the balance of the entire economic structure. They cause positive waste of economic power, by inducing misapplication of resources during the speculative phases of prosperity periods and precipitating idleness of resources during the subsequent depression. Nothing need be said of the incalculable human misery caused by prolonged depression; this is too well known to require emphasis. The broader impersonal consequences of the cycle are sufficiently serious to justify the statement that cyclical fluctuations are the chief cause of many of our most grievous social problems.

These cyclical fluctuations of business are features of an economic system centered around the getting and spending of money incomes, a system in which groups of specialized productive processes serve uncertain markets under the guidance of prospective pecuniary profits. The conclusion does not follow that an economic system of this type — a business economy — should be replaced by some other form of organization. Historically the vicissitudes of non-business economies have been even more violent and humanly disastrous, although they have been episodic in character rather than cyclical. The conclusion is, not to destroy the business system with its cyclical disruptions, but to seek measures that may control so far as possible the cyclical disturbances.

Control of cyclical changes: Conditions of the problem. Apparently an individualistic, competitive economic system, organized around prices, cannot automatically overcome the forces causing cyclical changes and adjust itself to these changes rapidly enough to avoid cumulative disequilibrium. On the other hand, the problem of applying external controls is by no means simple. There are two reasons for this.

In the first place, there is the question of what the objective of control should be. Since the crisis of 1929 the attention of most economists has been focused on the problem of bridging the gap between depression and revival, of finding methods of giving an impetus to business activity which would initiate the upward spiral of increased employment, output, and income. But there is also the problem of trying to restrain the fluctuations of business activity within narrower bounds. This implies the application of controls in the period of revival or during the early stages of prosperity in the hope of avoiding such an accumulation of internal strains as would make inevitable a lengthy and disastrous period of liquidation.

In the second place, there is not substantial agreement among economists as to the measures which are appropriate to achieve either of these objectives. This is not surprising in view of the fact that the two points in the analysis of the business cycle on which economists differ most widely have to do with the forces which bring about revival and those which turn prosperity into recession.

The measures of control which have been proposed are so numerous that we can only touch briefly on a few which seem promising. These proposals all encounter one common problem, since all require intelligent, disinterested, and courageous administration by some central authority. One's judgment as to the practical virtue of any proposal depends therefore not alone on the verdict as to the abstract soundness of the measure but quite as much on his confidence in the governmental body which is to administer it

Flexible prices. The suggestion has frequently been made that the control of cyclical fluctuations might properly take the negative form of securing greater flexibility in the economic system, especially with respect to the price structure. Complete flexibility in any system of human relations and humanly operated institutions is obviously impossible. But with a greater degree of flexibility of prices than now exists, cyclical variations would be less extreme; for it is the failure of prices to move in harmony which causes most of the trouble. When the price level is being shifted by uncontrollable forces, rigidity in any group of prices

will throw an increased burden of readjustment upon the more flexible areas. Even with more flexible prices cyclical fluctuations would still be serious, and there would always be the tendency for the economic system to move back again toward greater rigidity. The self-interest of groups and individuals leads them to resist price changes, including changes in wages, which are immediately unfavorable to themselves. These resistance policies take very many forms: commodity valorizations, coöperative marketing, industrial concentration, agreements within trade associations, the gentlemanly following of a leader so as not to "spoil markets," a widening of the scope of regulation of public utilities. But despite the difficulties which confront any government in its attempts to prevent monopoly and restraint on trade and to maintain truly competitive and flexible prices, there is every reason for making strenuous effort to preserve those conditions which promote price flexibility all along the line.

Control of wage rates. Another suggestion which has frequently been made is that some authority exercise direct control over wage rates. If in the early stages of revival wages could be raised sufficiently to neutralize that particular tendency toward increasing the margin of profit, some of the incentive toward rapid and speculative expansion of business would be removed. In further defense of this policy it is also argued that it would place increased purchasing power in the hands of labor and thus tend to maintain a better balance between consumption and investment. Such a plan has serious practical obstacles to face, for a rapid rise in wages in a period of revival would necessarily carry with it a stipulation that wages should be decreased in a period of falling prices, when the profit margin was declining, and this would invite the determined hostility of labor. It is similarly certain that employers would not submit without a struggle to an arrangement which would thus curtail their profits. This struggle might take the form of a reclassification of the more highly paid labor, in order to place them in a lower wage group, or a greater resort to the use of capital. Despite these practical difficulties, the proposal to secure greater flexibility in the price of labor by positive measures of control is deserving of serious consideration.

The rate of interest. For many years economists have been interested in the possibility of controlling the rate of investment and by this means of exercising some measure of control over the business cycle. This is an instrument of control which lies ready to hand, for in all capitalistic societies the rate of interest is regulated to a considerable degree by action of the central bank. A policy of high discount rates designed to restrict lending by banks is recognized as generally effective to discourage investment. It has seemed only natural to inquire whether or not a corresponding easy money policy, in the United States on the part of the federal reserve system let us say, would stimulate investment.

The rediscount rate, being the rate at which member banks may borrow from the federal reserve banks, has its effect upon the rates of interest charged by member banks to their customers. The direct effect, to be sure, is only upon a narrow group of interest rates. But there is available also the instrument of open-market purchases of government bonds, which either increase the excess reserves or diminish the indebtedness of the member banks to the federal reserve banks, thus putting strong pressure upon member banks to find other outlets for the investment of their funds. A central bank therefore, if it is sufficiently resolute, cannot only exercise a rather powerful and direct control over bank rates of interest, but can also exert pressure upon the banks which will affect long-term interest rates.

In view of these powers possessed by the central bank, it might seem that monetary control of the rate of investment should work effectively. The history of the last fifteen years however does not confirm this expectation, for reasons that are by now quite generally accepted. Although the interest rate does undoubtedly tend to affect the rate of investment in the manner just described, the influence exerted by the rate of interest does not appear to be sufficiently powerful to achieve the results expected of it. Particularly is it true that when a recession has once started, a lowering of the rate of interest is virtually powerless to prevent the recession from running its course. Low rates of interest, as they spread through the economic system, may do much to promote an up-turn if other factors are favorable. And the high rates of interest that prevail

during the late stages of a boom are frequently sufficient of themselves to dampen the rate of investment and thus to precipitate a depression. It may be concluded then that control of the rate of interest might be used with some success to prevent a revival from progressing too rapidly or a period of prosperity from turning into an inflationary boom but that such control certainly will not be effective to halt a depression when once started, though it might later assist in bringing about a revival.

Public works. In recent years many governments have intervened directly in the economic lives of their communities in the effort to increase the rate of spending. This has been done in the democratic countries, not by enforcing rules of conduct upon individual and private enterprise, but rather by entering directly into the function of investment through the medium of public works and relief financed by borrowing.

When a government spends money in excess of its receipts, the effect upon the economic system as a whole, so far as concerns the relationships which we are here investigating, is similar to that of an equal volume of private investment financed by borrowing, and the effects of such deficit spending may be traced out in precisely the same manner as the effects of private investment. An increase in government spending will lead to a simultaneous increase in the incomes of individuals or business men who receive the benefit of the spending. They in turn will be in position to increase their expenditures and thus increase the incomes of a second set of individuals. Thus the expenditure of, let us say, a billion dollars of borrowed money by government may produce an increase in aggregate gross income of several times that amount. The proponents of government spending argue that, if it is carried out on a sufficiently large scale, it may initiate a healthy recovery of private investment. They reason that an increase in aggregate income always stimulates investment and that deficit spending by the government, when its full effects upon income have worked themselves out, will initiate a normal recovery and thus permit the deficit spending itself to be halted.

It must be noted however that the similarity between private and government spending applies only to the act of spending the

money. When private enterprise borrows funds for investment, there is normally the expectation that the undertaking will produce an income sufficient to cover (1) all operating costs, (2) interest on the funds invested, and (3) amortization of the loan; *i.e.*, to accumulate by the time the loan matures a reserve fund equal to the amount borrowed. Public expenditure on the contrary often involves the government in debt without leading to the creation of income producing assets capable of yielding a surplus over costs for the payment of interest and amortizations. In this respect government deficit spending contrasts sharply with the investment of borrowed funds by a private enterprise. It is for this reason that private investment may continue indefinitely in an expanding economic system. Although it usually involves a steadily increasing total of private indebtedness, it also involves the creation of a larger and larger volume of capital capable of earning a profit for its owners and thus enabling them to carry the interest charges they have incurred. In contrast public deficit spending, if continued indefinitely, has the effect simply of increasing the interest burden that must be borne by the government which finances it.

Furthermore the favorable effects of government borrowing upon income may be offset by an unfavorable effect on business, if it results in raising interest rates and making capital difficult to obtain. Therefore the policy of deficit spending must, to be effective, be accompanied by an easy money policy on the part of the central bank. Again if government spending is carried out in such a way as to increase costs of production, and above all costs of investment for the private business man, then its net effect upon private investment may be unfavorable instead of favorable. Similarly if government funds are spent in the creation of agencies and facilities which compete with private business, the net effect of the spending may again be unfavorable to private investment. A final point, perhaps more important than any of the preceding, is that the continuation of a large government deficit over a period of years may have depressing effects upon business confidence. Business men envisage the prospect of currency inflation, which in the past has generally followed in the wake of deficit financing by the government; such inflation injects uncertainty into all business

conditions and gives to investment the character of gambling. Business men take alarm from the possibility of actual collapse of government finance and the unpredictable actions which government may take in that event. More obviously the business world is disturbed by the prospect of further additions to the burden of taxation, thus adding to business costs elements which cannot be predicted or intelligently provided for.

We may conclude that a program of deficit spending will for the time being increase the aggregate income and reduce the volume of unemployment, but that if it is to have the desired secondary effects upon private investment it must be carefully engineered so as to avoid unfavorable effects upon interest rates, costs, or the demand for the services of private business. Above all it must be financed in a manner that commends itself to the business world as sound. Because the results of government spending depend upon so many intangible factors, its ultimate effect will depend upon the way in which the program is carried out. Even though a public spending program be carried out in such way that it does not impair the confidence of business men, it will not suffice by itself to produce a period of prosperity unless substantial private investment opportunities appear. Public spending may be successful, if it merely initiates a revival for which the stage is already set. But if no substantial private investment opportunities appear, then the revival is not likely to last when government deficit spending is stopped. These conclusions are strongly supported by the experience of the United States during the last few years.

XL

LARGE-SCALE PRODUCTION AND COMBINATION

The growth of large-scale production. A consequence of the Industrial Revolution has been an increase in the size of the producing unit as measured by the amount of capital used, the number of workers employed, and the volume of output The table on the page following gives a statistical view of this growth for cotton manufacturing and for the boot and shoe industry in the United States during recent periods[1]

A study of these statistics reveals several interesting features. In the years from 1859 to 1889 there was a slight decline in the number of establishments devoted to cotton manufacturing, while for the period as a whole there was an increase of about fifty-five per cent More striking than the increase in the number of establishments is the fact that for the entire period there was an increase of 155 per cent in the average number of workers per establishment, an increase in the primary horsepower per establishment of over 800 per cent, and an increase in the average value added to the product by manufacturing per establishment of 334 per cent.

In the boot and shoe industry the situation is nearly parallel. In the half century following 1879 the number of workers per establishment increased by 150 per cent, the average primary horsepower per establishment by 1500 per cent, and the average value added to the product by manufacturing per establishment by 400 per cent. The number of establishments in the industry shows a significant decline, whereas for cotton manufacturing there is during the same period (1879 1925), a great increase Part of the explanation of

[1] Census of Manufactures, U S Bureau of the Census 1925 Derived from tables on pp 230 and 598 Values have been reduced to a common unit by the use of index numbers supplied in part by the National Industrial Conference Board, base year 1913

Reports of the U S Census are the source of the statistics presented in the pages immediately following

this difference is found in the fact that the cotton industry was on a factory basis at the opening of the second half of the nineteenth century, whereas shoe manufacturing was still in large measure a local industry carried on in a very small way. The application of factory methods to shoe manufacturing meant the replacing of a large number of small producing units by a smaller number of units operating on a larger scale.

Year	*Number of establishments**	*Average number of wage earners per establishment*	*Average primary horsepower per establishment*	*Average value added to product by manufacturing per establishment*
		COTTON MANUFACTURES		
1859	1,091	112	—†	$ 60,824
1869	956	142	153	50,947
1879	756‡	228	365	139,869
1889	905	212	513	150 527
1899	1,055	287	754	205,558
1909	1,324	287	979	200,411
1919	1,496	299	1,243	285,767
1925	1,638	286	1 392	263,999
		BOOTS AND SHOES (LEATHER)		
1879	1,959	57	6	$ 38 200
1889	2,082	64	15	58,946
1899	1,599	89	31	75 327
1909	1,343	138	64	126,980
1919	1,149	146	83	147,330
1925	1,460	142	96	191 156

* As a rule the term "establishment" represents a single plant or factory but in some cases it represents two or more plants which were operating under a common ownership

† Figure not available

‡ Figures exclude 249 mills, classed as "special mills," making hosiery, tapes, fancy fabrics, and mixed goods

If we were to examine the statistics showing the growth of other industries we would find differences in detail but the same general trend — a great increase in the average number of workers, the average amount of capital, and the average output per establishment. While the figures given are sufficient to show clearly the increase in the average size of the industrial plant, they do not demonstrate the significance of the very large plant as compared

with the plant of small or moderate size. This is brought out by the table below.

ESTABLISHMENTS GROUPED BY VALUE OF PRODUCT
ALL INDUSTRIES, 1929

Value of product	*Per cent of total number of establishments*	*Per cent of value of product to total*	*Per cent of number of wage earners to total*
$ 5,000 to $ 19,999	32 9	1.1	2 3
20,000 to 99,999	35.7	5 1	7 8
100,000 to 499,999	20 9	14 2	18.9
500,000 to 999,999	4 9	10 4	12 7
1,000,000 and over	5.6	69 2	58 2

The establishments of small size are seen to be most numerous; those producing between $5,000 and $100,000 of goods a year formed about 68 per cent of the total number of producing units, while those with an annual output of a million dollars or more constituted only 5.6 per cent of the total. This comparison would be even more marked were establishments with an output of less than $5,000 included in these census figures. The 1920 Census, which gives these figures, shows that for 1919 the number of establishments producing less than $100,000 of goods annually formed 79.6 per cent of the total number and that nearly 53 per cent of the total produced less than $20,000 of goods annually. Turning to the other columns in this table, we observe how important the large-scale establishments are in spite of their relatively small numbers. Plants with an output of less than $100,000 — over two thirds of the total number — employed but 10 per cent of the total number of workers and contributed but 6 per cent of all the goods produced; whereas the plants with an annual output of $1,000,000 and more employed 58 per cent of the workers and turned out 69 per cent of the total goods produced measured in terms of value.

The picture may be made more vivid if we approach it from a still different angle. In 1929 the establishments employing from one to five laborers comprised 45.4 per cent of the total number, whereas the establishments employing over 500 workers formed 1.3 per cent of the total; these large establishments however employed 37.8 per cent of the total wage earners in industrial plants,

as contrasted with the 3.2 per cent employed by the establishments in the former group. Establishments employing more than 1,000 laborers formed only 0.47 per cent of the total number of establishments, but they employed 24.4 per cent of the total number of workers. That this trend towards large-scale production is not abating is seen from the fact that from 1914 to 1929 the average number of employees in all industrial establishments increased from 39 to 42.

Two conclusions of considerable importance can be drawn from these figures. The first is that small-scale industry is still of significance in our national life, particularly to supply local markets. "Big business" has not entirely supplanted the small producer, even in the lines of industry where the advantages of large-scale production seem most marked. In steel works and rolling mills there were in 1929, out of a total of 486 establishments, but 194 which employed an average working force of over 500, and in the cotton goods industry only 253 out of a total of 1,281 establishments employed more than 500 workers.

At the same time it is true that the large units have been growing, both absolutely and relatively, at the expense of the smaller units. Production for national markets has been supplanting production for local or regional markets, and with this change has come the decline of the small unit and the substitution of the larger.

Conditions necessary for large-scale production. One of the first prerequisites for large-scale production is obviously a product which is in great demand, for unless a large output can be marketed profitably there is no scope for increased refinements in the productive apparatus. Another requirement is that the article be not one which is demanded because of the finish which hand labor may give it. It must be amenable to the machine process, capable of standardization, and one which does not entirely lose attractiveness in the eyes of the purchaser thereby. Many articles of the finer grades of clothing, of personal adornment, and of luxury can never be produced by large-scale methods.

Economies of large-scale production. Large-scale production, where it is found, exists for the simple reason that it is more efficient

than small-scale production, and by this we mean that the larger unit effects economies in production which make it possible to sell the good at a lower price. This not only extends the range of commodities over which the consumer has command and therefore enlarges his living, but it is advantageous in a more profound sense, for it implies the more effective utilization of the limited supplies of labor and capital. Let us consider some of the more important ways in which economies are realized.

Power and machinery. One great economy of large-scale production results from saving in the use of power-driven machinery, both in the use of power as well as in the use of machinery. The saving in power arises from the lower unit cost of installing and operating a power plant Though a large plant costs more to install than a small one, it costs proportionately less, and greater efficiency in the production of steam results Hence the cost of power to be attributed to each finished article is less

Machinery can be used for more of the processes in a large factory than in a small one, for if a machine is to be used to advantage it must be used more or less continuously. In determining the unit cost of production by machine one must take into account the yearly interest on the investment and also a sum large enough to cover repairs and the eventual replacement of the machine, as well as the power and the labor necessary to run it If the machine is used only a part of the time the cost which must be borne by each unit of the good undergoing the particular process may make use of the machine more expensive than hand labor. A large factory is able to employ machines which save only a fraction of a cent over the cost of hand labor but which are distinctly worth while on account of the large volume of production This is one way of stating the proposition that large-scale production permits of the greater application of the principle of division of labor There can be greater specialization by workers, and processes can be more finely divided and therefore more easily subjected to machine methods.

By-products. Another economy lies in the utilization of by-products The meat-packing industry offers one of the most illuminating examples of this. Swift & Company give the following summary of the products derived from a steer:

Percentage of weight of finished products to weight of live steer

Beef	54.3%
By-products:	
Hide	5.9
Fats	2.2
Head	2.2
Feet	1.1
Blood	.7
Casings	.8
Miscellaneous	3.2
Valueless	10.1
Shrinkage	19.5
	100.0

A list of the finished goods which are made in whole or in part by Swift & Company from these by-products includes the following: glue, fertilizer, oleomargarine, lard compounds, soaps, glycerine, leather dressing, animal feed, buttons, combs, gelatine, ice cream, candy, pharmaceutical products, snuff containers, gold beaters' skins, and sausage containers. In the early days the materials from which these goods are now made had to be carted from the city and buried; so far from being a gain, they were the occasion of considerable expense.

Only when the quantity of by-products is large is it worth while for the company concerned to refine them. Otherwise they are thrown out as waste or perhaps sold to a subsidiary industry organized to handle the by-products of a number of similar industries. The large industry however is likely to make a greater profit from refining its by-products itself than the small producer can get by selling them to another concern. Unless competition to secure the by-products sets in, the small producer finds it necessary to get rid of them as best he can, and he is usually glad to take any price offered in order to avoid the expense of disposing of them as waste.

The economy which results is plain; any net return obtained from the by-products diminishes by so much the price which must be obtained for the principal product in order to cover the expenses of production. The meat packers for example often pay more for live cattle and hogs than they receive for the dressed meat, but they make up the difference in the return from by-products.

Technical experiment and research. A large company is better able to conduct experiments than a small company. Experimentation is a risky and expensive affair. Many trials are necessary before a device or a process is developed to the point where it is of practical use. For a small concern the financial burden may be too heavy to carry, whereas the company with greater resources can carry the load through the lean years, trusting that in the long run the improvement will more than pay for itself.

In the making of pig iron chemists were not employed in this country to analyze the ore and the materials used with it in smelting until after 1870. The blast-furnace manager was supposed to be able to diagnose the condition of his furnaces by instinct. The result was that inferior grades of ore oftentimes enjoyed a better reputation than those of better quality, since a foreman who had evolved a mixture which suited low grade ore often rejected ore with a higher iron content simply because the mixture did not work. Mr. Carnegie was one of the first to employ a chemist, and he reaped an immediate reward in that he was able to secure ores at low price which were found on analysis to contain unusually high percentages of pure iron. In his *Autobiography* he says that years after he had taken a chemist as his guide some of the proprietors of other furnaces claimed that they could not afford that luxury.

The present-day importance of research in the fields of production, marketing, and pure science is well illustrated in the case of the United States Steel Corporation. In 1937, according to a statement of the Chairman of the Board of Directors, the corporation had 174 laboratory departments and spent $8,400,000 on metallurgy and research. A survey conducted by the Mellon Institute for Industrial Research showed that in 1937 about $100,000,000 was spent for industrial research in the United States.

Economy in administration. The cost of administration and supervision is not much greater for a large factory than for a small one, certainly not proportionally greater. Within limits the head of a department can generally take care of a large department with about the same effectiveness as of a small department. The chief additional expense consists in a larger clerical force. Similarly the cost of buying raw materials and selling finished goods may be

decreased per unit. Selling in large quantities means a real saving to the seller in packing, shipping, and sales costs, and he can therefore afford to quote a lower price to one who buys on a large scale. In selling the finished good there is better opportunity to coördinate the activities of salesmen, thus covering a given territory without overlapping and with greater effectiveness, while national advertising can be used with greater profit.

An interesting statistical study of the costs of production of large and small plants is presented in the report of the Federal Trade Commission on the canned salmon industry. In 1916 the average unit cost of production for large plants was $3.491 and for small plants $4.111; in 1917 the figures were $4.302 and $5.681 for the large and small plants respectively.

The field of large-scale production and its limits. Our attention thus far has been devoted exclusively to manufacturing, the field in which large-scale production obviously finds its greatest opportunities and from which may be drawn the most illuminating illustrations of its principles. Many other lines of industry have however proved themselves amenable to large-scale production. In transportation, banking, and marketing the size of the organization has greatly increased in the last century. Doubtless this increase in size has come partly in response to the growth in the size of manufacturing establishments, while at the same time it has been one of the causes of that growth. In any event concentration of manufacturing in large plants requires more effective service than could be afforded by the transportation, banking, or marketing system of a century ago.

The one great industry which has not proved a profitable field for the application of large-scale methods is agriculture. In the tropics we do find large rubber or tea plantations, and even in the United States there are of course some examples of farms where hundreds of workers are employed, but they are the exception and not the rule. The long hours of work on the farm are not attractive to most laborers, and the difficulties of supervision on the farm are so great that unless the laborer is working with or near the owner he is tempted to loaf on the job. In a factory individual supervision by the foremen and automatic checks such as time clocks, machin-

ery with controlled speed, and the piece-rate system of wage payments do away with this difficulty to a great extent, but on the farm the spur of individual gain seems necessary to secure continuous effort. Hence we find that in most cases which do approach large-scale farming the tenant system is used. The owner of the farm provides materials and tools, and the farm laborer receives a certain fraction of the output, thus giving him a direct incentive to increase his efforts.

Large-scale production has its limits, even in the fields most favorable to it. Probably for each industry there is a plant of optimum size, beyond which it is unwise to attempt to go. This is due in part to the physical awkwardness of handling a large bulk of goods under one roof, but a still greater difficulty is the weakness in management which may appear. Since this weakness is common also to combinations of several plants under one management, its further discussion may better be postponed until after our investigation of the problem of industrial combination.

Combination. Rarely does large-scale production give rise to problems of vital concern to society or raise the question of regulation in the public interest. Sentimentalists deplore the elimination of many small shops and the passing of the independent craftsman, but while we may have lost thereby a distinctive and virile element in our economic life, few would seriously suggest that we sacrifice the greater gains which have come from mass production in order to try to recapture the spirit of an earlier day.

This however is not always true of the *combination*, which we may define as *a union of several industrial units under a common management*. There are indeed serious problems of contemporary interest and importance raised by the great corporations, the combinations, and the combinations of combinations. These will engage our attention in the remainder of this chapter and in several of the succeeding ones.

The large corporation. We have become so used to the rôle played by the great corporations in our everyday life that, paradoxically perhaps, it is sometimes rather surprising to discover how large they loom in the economic picture. According to one estimate the two hundred largest non-banking corporations in the

United States had at the beginning of 1930 total assets of $81,074,-000,000; the assets of all the non-banking corporations, something over 300,000 in number, amounted to $165,000,000,000. Almost half of the assets of all non-banking corporations were in the control of less than one per cent of the total number of such corporations. The same authorities also estimate that these corporations owned thirty-eight per cent of all business wealth, other than banking, and twenty-two per cent of all national wealth.[1]

Another estimate, for the year 1933, shows that 375 non-financial corporations, or 0.13 per cent of the total number of 287,575, owned 56 per cent of the total assets of these corporations. The 594 largest corporations in all fields, or 0.15 per cent of the total number, owned approximately 53 per cent of total corporate assets.[2]

A somewhat different picture of the importance of large corporations is given in the following table, which shows the number of corporations in selected lines having in 1929 an income of $1,000,000 or more and the percentage of total net income in the given line which accrued to these larger corporations.[3]

Concentration of Corporate Control

Industry	*Number of corporations with net income of $1,000,000 up*	*Net income (millions)*	*Per cent of all net income*
Manufactures	627	$3,338	64 0
Mining	65	278	84 6
Public utilities	230	1,805	86.0
Trade	93	316	27 5
Service	31	108	34 4
Finance	283	1,048	47 7
Total	1,329	6,893	60 5

In recent years the large corporations have shown a growth which is much more rapid than that of the small or medium sized corporations. From 1907 to 1927 the assets of the two hundred largest non-financial corporations increased more than twice as rapidly as the assets of other non-financial corporations; they rein-

[1] For a further discussion of this topic and for an illuminating analysis of the social implications of the position of the modern corporation the student is referred to Berle and Means, *The Modern Corporation and Private Property*, New York, 1933, from which these figures are taken

[2] *Big Business Its Growth and Its Place* Twentieth Century Fund, Inc., New York, 1937, p 5.

[3] Lewis Corey, *The Decline of American Capitalism*, New York, 1934, p. 387

vested a larger portion of their earnings; they secured a larger share of new capital by the flotation of securities; and mergers were more common and more significant among them.[1]

The formation of combinations. Many a corporation has become a combination quite simply and naturally when it was found profitable to build a new plant from earnings or from funds secured by the floating of additional securities. Some have grown by the consolidation of two or more existing organizations; one company has purchased the property of another company and then completely merged the two into a larger single company. The American Tobacco Company, formed in 1890, exchanged its stock for the plants, business, brands, and good-will of five companies, and the American Sugar Refining Company, organized in 1891, is another example among the many that might be cited.

The most common, the most effective, and the least expensive method of building up a large combination is however through the holding company. A corporation which buys the controlling shares of stock in another company, either with its own stock or with cash, becomes in effect a holding company. Frequently a new corporation is chartered, as in the case of the United States Steel Corporation, to buy up a controlling number of shares in the companies which are to be combined. Little capital is required to start a holding company, since the shares of stock of the member companies may usually be obtained in exchange for the stock or bonds of the holding company. Unity of policy may be obtained by the election of the proper men as directors in the constituent companies.

The development of the holding company was delayed for some time by the failure of state legislatures to grant to corporations the power to hold the stocks of other corporations. In 1889 and 1893 however New Jersey amended its corporation laws so as to permit corporations to hold stocks in other companies and to exercise the rights of owners, and the fear lest New Jersey become the home of all large corporations led some of the other states, notably Maine, New York, and Delaware, to liberalize their laws in the same fashion.

[1] Gardner C. Means, "The Growth in the Relative Importance of the Large Corporation in American Economic Life," *American Economic Review*, October, 1931, pp. 10–42.

Motives leading to the formation of combinations. In considerable measure combination has been resorted to as a means of securing substantial control of a particular field of production in order to obtain monopoly profits. Two other motives however have been prominent. Of these the first has been to make possible speculative profits through the issuance of large volumes of watered stock; the second has been to obtain even greater economies in production than were possible simply through large-scale production.

We have therefore three distinct problems. The first, and in some aspects the most important, concerns itself with the growth of monopolies — the so-called "trusts" — and their effect on society and the methods by which they may be regulated in the public interest. The second involves methods used in marketing securities and the means by which the operations of investment bankers, underwriters, and the security exchanges may be controlled in order to prevent abuses. The third problem, to which the remainder of this chapter will be devoted, centres around the possible social advantages of combinations as producing units. Are there economies in production possible to a number of large plants under single management which cannot be obtained by a single large producing unit? If there are such economies and if they are of considerable significance, this fact will have an important bearing on the first problem.

Horizontal combinations: Definition. We may distinguish two types of combination, *horizontal* and *vertical*. As the name suggests, a horizontal combination is one which unites under a single management several plants producing the same product or products. It is a type which is quite common in the United States, being exemplified by many of the "trusts" which exist or have existed in the past, such as the American Sugar Refining Company, the United Shoe Machinery Company, the American Tobacco Company, and others.

Economy in buying. Such a combination may possibly secure a saving in the purchase of its raw materials. Buying in much larger quantities than the single plant, it may be able to afford the services of a peculiarly shrewd buyer, or it may approach the producer directly and eliminate the middleman's profits, or it may secure its

goods from a wider productive area and thus obtain its raw materials more cheaply. But there is certainly some limit to the savings in this line, and it is to be doubted if they are great enough to be of real consequence. Of course if the combination controls a large part of the total production it may use its power to depress the price of its raw material and cut the profits of the producers. But this does not constitute a saving in the social sense, as it merely gives to one class what is taken from another, nor does it give the combination an advantage over its competitors.

Economy in selling. In the sale of its products there is a possibility of saving. A reduction in the number of salesmen and in the amount of advertising and greater effectiveness in the use of both sales media may lower the sales cost. The tobacco trust, which had been formed by the combination of several large producing units, was dissolved by a court decree in 1911, whereupon the business was turned over to a number of separate companies. "The subdivision of the business provided for in the dissolution decree led to a duplication of selling organization and an increased overhead expense; and the result was a general increase in selling costs. . . . In every branch, except flat plug and Turkish cigarettes, the selling costs showed an increase. . . . The advertising expenditures of the successor companies also greatly increased as compared with the expenses of the trust. The advertising expenses of the trust in 1910 in all branches except cigars were $10,895,132, while those of the successor companies in 1913 amounted to $23,623,564, or more than double."[1]

Economies in operation. The greatest field for economies would appear to lie in the operation of the plants. Each plant can presumably secure maximum efficiency by producing a certain number of articles. If forced above that point or below it, the unit cost of goods is increased. With several plants operating under unified control it is possible so to arrange things that all the slack in production will fall on one or more plants, leaving the others running as usual. The American Sugar Refining Company is said to have used its Brooklyn refinery for this purpose, slowing down operations there as demand slackened and speeding up in response to in-

[1] Eliot Jones, *The Trust Problem in the United States*, 1921, pp 149-150.

creased demand. The old Whiskey Trust, by adopting a similar policy with respect to its distilleries, was able to effect a large saving in the cost of manufacture. In some industries the risks and uncertainties involved in production may be reduced if plants are located in different parts of the country. A labor strike may not affect all of the plants at the same time, and a tie-up of the transportation facilities of all the plants at one time is unlikely, so that if the closing of some is forced the others can be operated more intensively. In the salmon canning industry the size of the run of fish is a decisive factor, as a small run of fish may entail heavy losses. In 1917 a company operating six plants lost $115,000 on one plant, but made a net profit on the six plants of over $1,000,000. The owner of a single plant is liable to have heavy losses one year and large profits another, and the stabilization of profits which comes from enlarging the area of operations is of considerable advantage.

The devices or patented articles that have proved most successful can be made the common property of all the operating units, heightening the efficiency of each. With several plants it is possible to start competition between them, thus enlisting the enthusiasm of one manager to excel the others. Cost statistics may be available for all the plants, so that if one falls down in some respect the trouble can quickly be diagnosed and remedied. It is claimed that more efficient executives can be secured because of the wider field from which to choose. The inter-plant competition, it is asserted, will disclose men who are most capable of handling the more difficult tasks of coördination and centralization.

Specialization may be carried to greater lengths by a combination than by a single producing plant. A combination which makes a variety of products is able to specialize in factories, devoting one factory to the making of a particular article, rather than allowing all factories to make all articles, and in this way the utmost advantages of specialization are attained.

Saving in freight costs. The savings on cross freights are sometimes cited as proving the superiority of the combination. If under separate management there are two plants, one in New York and the other in Chicago, each may sell goods in the other's district and

incur freight charges unnecessarily great. Under combination the good will be sent from the nearest point. But the importance of this saving, while real in many cases, can be greatly exaggerated, for if the good be bulky and of low value the distance it can be shipped is limited by the cost of freight, whereas if it be a small and expensive object such as a watch, freight charges form such a small portion of its value as to be of little moment.

Vertical combinations: Definition. Vertical combination involves the gathering under one management of the various processes necessary to turn out a finished good. It is sometimes called *integration* of industry and is the opposite of that process of differentiation and specialization which has been characteristic of the factory system. In some industries there has always been a certain amount of integration. Thus in the United States spinning and weaving have commonly been undertaken by the same factory, although in England they are separate industries. In the steel industry of the United States can be found the most impressive illustrations of vertical combinations, although horizontal combinations also have been common.

Technical advantages. There are real technical advantages in combining thus the different stages of manufacture of steel products, advantages which result in saving of fuel for reheating the metal, saving of time and labor in moving or handling material, and more complete utilization of by-products, such as blast-furnace gas. Further advantages may be obtained through the control of sources of iron and coal. In the early history of the steel industry it was not considered essential to own iron fields, but as concentration in the ownership of the known and proved fields developed, the unpleasant prospect of being forced to pay large profits to the producers of iron ore had to be faced. The alternative was ownership or control of iron reserves. This held true of coal and coke as well. The steel industry differs from some others in that each unit not only may use the products of another as raw material but may pass its own finished or semi-finished products on to serve as raw materials for the others. Coal and iron are essential for the blast furnaces, while the ultimate products of the blast furnaces are requisite in the mining and transportation of coal and iron.

The United States Steel Corporation. The extent of the process of integration in the United States can best be pictured by a brief summary of the organization and scope of operations of the United States Steel Corporation, one of the largest corporations in the United States, incorporated in 1901 with an authorized capital stock of $1,100,000,000. The corporation secured control of a number of other corporations by an exchange of stocks and bonds, and at the outset it controlled eleven corporations: (1) The American Bridge Company, fabricating bridge material and structural steel, (2) The American Tin Plate Company, (3) The American Steel Hoop Company, producing chiefly iron and steel bars, hoops and bands, and cotton ties, (4) The American Sheet Steel Company, engaged in the manufacture of practically a single product—sheets, (5) The American Steel and Wire Company of New Jersey, manufacturing wire nails, plain wire, barbed wire, and wire fencing, (6) The National Tube Company, producing iron and steel wrought tubing, (7) The Shelby Steel Tube Company, a consolidation of practically all the important companies making drawn or seamless tubing, (8) The Lake Superior Consolidated Iron Mines, which was an ore producing and transportation company exclusively, controlling the Duluth, Missabe & Northern Railroad (this railroad had a controlling interest in the largest fleet of Lake ore carriers, the Bessemer Steamship Company, which the United States Steel Corporation later purchased), (9) The National Steel Company, confined largely to the production of semi-finished steel, such as billets, sheet bars, and tin-plate bars, (10) The Federal Steel Company, itself a combination controlling coal, iron ore, and transportation, and turning out about the same line of products as (11) The Carnegie Steel Company, the most important of the group. The Carnegie Steel Company was itself the result of merger and reorganization and exemplified a high degree of integration. It had assured itself of extensive reserves of desirable ore, coal, and coke, had made a beginning in controlling the transportation of its own ore on the Great Lakes, and had acquired several railroad lines to ensure the proper and cheap handling of its raw materials and finished products. Its chief business was the manufacture of heavy steel products, such as steel rails, plates, structural steel, and crude steel

which served as the primary material for other concerns, paying little attention to the manufacture of finished goods.

It is a fact of interest that almost all the companies which were united to form the United States Steel Corporation were themselves combinations. Some of them represented merely the horizontal type of combination, such as the American Steel Hoop Company, which was a consolidation of nine concerns formed with the desire to limit competition and secure larger profits, while some already showed signs of a fairly complete integration, as the Federal Steel Company and the Carnegie Company. The United States Steel Company was not exclusively a vertical combination, for several of the constituent companies produced the same articles.

The result of the consolidation was the creation of a giant corporation, which was practically self-sufficient so far as materials were concerned and which embraced the manufacture of practically every known article in steel and iron. Its history since that time is one into which we cannot go, but we may say that in brief it consists of an extension of the processes already noted — the acquisition of more mines, furnaces, plants, steamships, railroads, and foreign selling agencies.

The question which immediately presents itself in this connection is whether the Steel Corporation has justified itself from the point of view of social saving. There is no doubt that the Steel Corporation has been a profitable venture; there may be doubt however as to whether the profits had their origin in economies of operation. In many instances of course the company avoided the payment of profits to others by the operation of a coal mine or a railroad itself, but so far as it merely took profits which would otherwise have gone to others this does not represent economy from the standpoint of society. So far as the economies to be obtained from the integration of manufacturing processes were concerned, they were probably already obtained to the full degree by some of the larger and better equipped among the constituent companies. The chief economies possible were from the more economical subdivision of the business among the plants, utilizing only the most efficient plants and dividing the work among the plants according to the nearness of the market. One writer in discussing this question states that

"The conclusion would seem to be that the Corporation, by virtue of its ownership of the cream of the ore and coking coal lands and of the iron ore railroads (not to mention its financial connections), had an advantage over its competitors but that this advantage did not demonstrate the superior economy of the trust form of organization." [1]

Combinations in Germany. For many years German economic organization has been marked by an increasing number of cartels or combinations. The law in Germany has not been so hostile to their formation as in the United States, nor public opinion so apprehensive of the dangers of monopoly. A typical cartel, embracing practically all the producers in a given line of industry, was able, by assigning quotas of production or by maintaining prices, to eliminate the competition which had existed previously and to achieve a stability of operations which is not likely to be secured under competitive conditions. If a cartel were engaged in foreign trade, it frequently pursued the policy of maintaining a higher price in the domestic market than in the foreign market. This it justified on the ground that the domestic prices would necessarily be still higher were it not for the outlet which the foreign market afforded. In the years immediately following the World War the number of cartels showed a decline, but the cartel continued as a significant feature in the industrial organization.

International cartels. In recent years certain combinations embracing the producers in several European countries have been formed, of which perhaps the most conspicuous example is the Continental Steel Entente, formed voluntarily by the majority of the steel producers on the Continent. The primary purpose of these cartels is to limit production by assigning production quotas to the producers of each country, thus reducing competition in the world markets and making it easier to keep world prices at a remunerative level. While the continental viewpoint that stabilization of industry is more wholesome than intense competition has led both governments and peoples to view these organizations with approval, few have operated successfully during the years of depression. In periods of increasing demand and of rising prices an international

[1] Eliot Jones, *The Trust Problem in the United States*, 1921, p. 219, *note*.

cartel may be reasonably successful, despite the incessant bickering over quotas, but in a period of falling prices it tends either to fall to pieces or to exert an unwholesome effect on the general economic situation by maintaining its prices at an artificially high level at a time when prices as a whole are falling.

Conclusions. It cannot be doubted that both vertical and horizontal combinations offer possibility of economies over and above those which can be obtained by the single large manufacturing plant. But it is exceedingly difficult to determine the extent to which these economies have been realized in actual fact. Combination and monopoly power are always so closely associated that it is not possible always to attribute the success of any given combination to the appropriate factor. It is however probably a fair statement of the case to say that where a combination has been conspicuously successful the profits have arisen in large part because of a monopolistic control of production which has enabled the combination to control the output and raise the price to the public. When monopoly is absent economies in the field of marketing are probably a more important factor than economies in production in the narrower sense.

As a matter of fact the success of the large combinations of the United States has not been striking, if success be measured in terms of profits. Not only have profits been smaller than anticipated by the promoters of the consolidations, but in many instances they have been smaller than the aggregate profits of the constituent companies in the period before consolidation. Despite the economies which combination might have been expected to bring, the United States Steel Corporation during its first year of operation earned but 78 per cent of the amount earned by the member companies in the period immediately preceding the formation of the combination, and for the first ten years of its life its average earnings were somewhat less than its earnings during the first year. Professor Dewing has made a comparison of the earnings of a number of representative combinations with the earnings of the companies which were united and has found for the group as a whole that the earnings of the first year, as well as the average earnings for the first ten years of the existence of the combinations, were

materially below the level of the previous earnings of the member companies.[1]

The failure of many combinations to reflect economies of operation and management in their earnings seems to indicate that fundamental defects or weaknesses are consequences of size of organization. A combination may realize some of the economies of large-scale management only to have these offset or more than offset by losses traceable to other factors. There are many losses to which a huge combination is peculiarly susceptible, but undoubtedly the chief weakness which organizations of tremendous size (including both large unit plants and combinations) have to face is the difficulty of finding men of the calibre required to direct the work. It is possible to find many an organization which seems to work with effortless, machine-like precision, but examination will usually disclose a dynamic personality behind the organization, supplying the initiative, making the important decisions, and inspiring his colleagues with enthusiasm and confidence. The founder of almost any of our large industries has been such a man. His personal fortunes have been closely tied up with the fortunes of the business, and not the least of his incentives has been the desire to leave his business as a monument to himself. When the reins of management pass from his hands into those of another, the same combination of ability, energy, and audacity is rarely found. When a leader of mediocre ability is in charge of a large business, lack of pressure for efficiency develops all along the line, and all employees are apt to have a feeling of security in the size and financial strength of the concern.

Business ability of the highest order is required by all large-scale production; it is essential for securing the economies of combination, but it is not sufficiently plentiful to give assurance that large combinations will be able by more efficient production to undersell the large single units. This conclusion does not however give warrant for the inference that the day of the large combination is over. On the contrary many indications point to a continued growth of the largest combinations. As pointed out earlier in the

[1] A. S Dewing, *The Financial Policy of Corporations*, 1934, vol. iv, Appendix. For other recent studies see *Recent Economic Changes*, 1929, vol. i, pp. 179–218.

chapter, the acquisition of monopoly power or the elimination of intense competition or the opportunity to secure profits from security issues may be sufficient motives for management to try to expand the scope of corporate activities, even though the result may be in opposition to the interests of society as a whole.

XLI

MONOPOLISTIC INDUSTRY AND ITS CONTROL

The nature of monopoly. Monopoly, as the term is used in economics, is the exclusive control of the supply of an economic good. The essence of monopoly is therefore the power lodged in the hands of someone to raise the price of the good which he offers for sale or restrict the quantity placed on the market. In the absence of legal regulation, monopoly price is usually higher than a competitive price would be, and monopoly profits are normally larger than the normal profits necessary to induce the production and sale of these goods.[1] In this sense therefore monopoly means the exploitation of consumers for the benefit of producers.

Capitalistic monopoly. Monopoly appears in many different forms. A city may create a monopoly by granting a railway company an exclusive franchise to lay tracks in the city streets, or a government may reserve to itself the exclusive right to manufacture and sell tobacco or liquor. The owner of the sole source of supply of some commodity or service has what is called a natural monopoly, in contrast to the legal monopolies created by the state. In this chapter however we shall be concerned with the so-called *capitalistic monopoly*, for it is this type which gives rise to the most serious public problems. Capitalistic monopoly rests on no exclusive control of the raw materials of industry, nor yet on any special privilege conferred upon it, although these may be present and may facilitate its growth. It arises rather out of the highly capitalistic method of production which characterizes modern industry and makes it possible for a concern both to acquire and to maintain its position of monopoly, not by virtue of superior productive ability, but because of more ample financial resources.

The capitalistic monopoly is a by-product both of large-scale production and of the corporate form of organization. When production is on a small scale, the number of actual or potential

[1] See Chapter XIV.

competitors is so great as to impede the formation of monopoly. Even if under these conditions a monopoly be formed, the capital outlay necessary to enable anyone to challenge the position of the monopoly is so small that its life is likely to be short. In certain lines of manufacturing however the superiority of the large producing unit is such that the small producers have to a considerable extent been eliminated, thus leaving the field mainly to a smaller number of large-scale units; this condition simplifies the problem of effecting a monopolistic combination.

In a still more significant way large-scale production has tended in the direction of monopoly. Such production generally means a relatively heavy permanent investment. This not only tends to discourage potential competitors, but it has a determining influence on the type of competition among those already in the field. Cut-throat competition is apt to develop in any industry where the fixed costs are heavy in relation to the variable costs, and the effects of competition of this type are so ruinous that even the threat of a price war may cause a scramble of competitors to seek shelter in combination.

The reader will doubtless recognize that complete monopoly — the exclusive control of supply — is seldom if ever attained by mere capitalistic monopoly without the aid of some other contributing cause. It is virtually certain that some other sellers will always be present in the market. However the serious public problems of monopoly do not wait upon the complete consummation of capitalistic monopoly but appear as soon as one concern has progressed so far in that direction as to be able to shape its selling policy without regard to its small rivals and so to wield a predominant influence in the control of price. Indeed the problem may arise even earlier than this; that is, as soon as there is evident a clear tendency toward monopoly of the capitalistic sort. Actually there are, outside the public utility field, very few real monopolies in the sense in which we define that term. The subject matter of this chapter is accordingly, not industrial monopoly in the strict sense, but monopolistic industry, or industry that approaches a state of monopoly.

Monopolistic industry: The problem. Popular support for governmental regulation of monopolies in the United States undoubt-

edly comes from a well-justified feeling that monopoly profits are representative not of any social service for which payment is due but of economic power which can take from the many and give to the few.

Valid as this reason is for public suspicion of monopolies, there is an even more powerful argument for regulation in the fact that the extension of monopoly control over any considerable portion of economic life would destroy the basis of private capitalism. We have observed in an earlier chapter the importance of competitive price as a regulator of economic life. We look to it to determine the distribution of the social income, to direct the flow of capital, to select the goods to be produced, to determine the quantities of each which shall be produced, and in other ways to promote the effective functioning of the diverse elements of our economic structure. At best it is not a perfect regulator, for adjustments do not take place immediately or automatically, and many maladjustments may require conscious control. In the case of public utilities, for example, competition usually develops into monopoly, and conscious control is requisite in setting rates which will give a fair return only on invested capital and laying down minimum standards of service. But should monopoly conditions hold sway over industry as a whole or be proceeding to that end, there could be no alternative to complete governmental control over economic life, for there is no reason to believe that monopoly price has any virtue as a regulator of economic life. It cannot be emphasized too strongly that the maintenance of truly competitive conditions is the prerequisite of economic liberty. Regulation to secure this should be viewed, not as an infringement of individual liberty, but as its preserver.

Historical. The trust movement. The term "trust," although it has another and technical meaning to be considered hereafter, has in common language come to be identified with the large industrial enterprise of monopolistic tendency, and we shall find it convenient to use the term with that significance. The reader must be on his guard however against the looseness of popular usage. He will not confuse mere size with monopoly power. Although size may be an indication of monopoly power, some of the very largest industrial

concerns are in active competition with rivals, perhaps equally large. Nor must combination be confused with monopoly. Though combination may frequently be a step toward monopoly, and though some combinations undoubtedly have and exercise monopoly power, cases of large combinations in keen price competition with each other, as for example the Carnegie Steel Company and the Federal Steel Company before the formation of the United States Steel Corporation, are not uncommon.

In the United States the movement toward combination for the sake of monopoly power — "the trust movement" — dates from the late seventies; in the eighties it gathered momentum, and in the nineties it attained its greatest proportions. Some indication of the extent of the movement can be gathered from the statistics of combinations formed. In the period from 1887 to 1897 eighty-six combinations were formed, in 1898 1900 one hundred and forty-nine, and one hundred and twenty-seven in the years 1901 to 1903. Relatively few of the combinations formed have been "trusts," but from 1898 to 1900 a larger portion of the combinations succeeded in acquiring monopoly powers in their respective fields than in any other period before or since. After 1903 the movement slowed down, partly because of the difficulty experienced in marketing the huge volume of securities issued at the time of the formation of each new combination, partly on account of adverse court decisions in anti-trust cases, and partly because of the fact that trusts had already been formed in most of the lines where monopoly was at all possible and indeed in some lines where monopoly power could not be retained.

With the close of the World War the formation of combinations again proceeded at a rapid pace. During the years 1919–1928, 1,268 combinations were formed in manufacturing and mining, involving the union of 4,135 and the disappearance of 5,991 separate concerns, while in the field of public utilities 3,744 public utility companies disappeared in the years 1919–1927 through acquisition by other companies. With the onset of the depression in 1929 the formation of combinations decreased abruptly.[1]

[1] *Big Business Its Growth and Its Place* Twentieth Century Fund, Inc., New York, 1937, p. 32.

How many of the combinations formed in these recent years are possessed of monopoly power is a question which cannot be answered, but in the opinion of competent observers the number is not insignificant.[1] The investigation authorized by Congress in 1938 may throw light on this question.

Devices restrictive of competition: The pool. Among the many devices used at one time or another to secure unity of action among competing producers, the pool is the oldest and probably the most frequently used. The term designates any sort of agreement between independent producers whereby their freedom of action in respect of certain matters is surrendered to a representative committee or to the group acting as a whole. Agreements to maintain prices, to divide profits, to restrict output, to divide the selling field among themselves, to provide for the common use of all patents, and so on, are examples of pooling agreements common to many lines of industry.

Pools have the advantage of retaining the identities and, so far as the methods of production are concerned, the independence of the member companies; furthermore they can be kept secret, and for that reason they are resorted to when other methods seem inexpedient. But while some of them have been quite successful, they have generally not been continuous, stable, or long-lived. The independence of each member is bound to bring about a conflict of interest. The restriction of output, which is necessary to secure higher prices, means of course rising costs of manufacture, and some black sheep among the members may succumb to the temptation to keep costs down by manufacturing more than the amount allotted to him and disposing of it without the knowledge of his partners. Furthermore the pooling agreement has never had any standing in the eyes of the law. Even in the absence of legislation making it positively illegal, any agreement which unduly restrains trade is, according to the common law, void and unenforceable. While the common law exacts no penalty for making such an agreement, the authority of the courts cannot be invoked to compel the parties to the pool to live up to their agreement

[1] Jenks and Clark, *The Trust Problem*, 1929. Cf. also W. L. Thorpe, "Persistence of the Merger Movement," *American Economic Review*, Supplement March. 1932.

or to pay damages to the other partners for breaking the agreement.

The trust. The instability of the pooling agreement led to a search for some device which would be stable and at the same time legal. The *trust*, created first in 1879 by the Standard Oil interests, seemed to meet these requirements. The trust, as technically defined,[1] consists of the creation of a board of trustees, to which are transferred the shares of stock in the several companies, together with the assignment of voting power. In return the former holders of the stock receive trust certificates, which give the holders a proportionate share of the joint net earnings of all the stocks held in trust. Since the trustees hold a controlling number of shares of stock in each company, they are able to elect the several boards of directors and thus to attain uniformity of policy and action.

Forty separate oil companies were thus brought together by the Standard Oil interests. The plan commended itself to others and was immediately adopted in the cottonseed and linseed oil business, by the whisky distillers, by the sugar refiners, and by others. But the formation of these trusts precipitated an anti-trust movement in the years around 1890. Various states passed anti-trust laws and attacked the trusts in the courts, and the passage of the Sherman Anti-Trust Act in 1890 indicated the awakening of the interest of the national government in questions of monopoly. In New York the North River Refining Company, controlled by the Sugar Trust, was attacked under the common law on the ground that it had exceeded its corporate powers in entrusting the direction of its affairs to an irresponsible board and that the trust was essentially a partnership, into which a New York corporation could not legally enter. As a result of this action the charter was revoked and the corporate existence of the company terminated. In Ohio the Attorney General brought action against the Standard Oil Company of Ohio, because of its connection with the trust, and succeeded in forcing the severance of the union. The effect of these

[1] It will be noted that the term "trust" is used in two senses From a technical viewpoint a trust is the form of combination here described, but the use of the term to describe any monopolistic combination whether in the form of trust or otherwise is too deeply imbedded in our language and thought to be rejected.

decisions was to make the trust formed to restrain trade illegal under the common law.

The holding company. The unfriendly reception accorded the trust led to the utilization of two other devices, the holding company and the consolidation or merger of competing companies, both of which have been described in the immediately preceding chapter. The holding company has certain advantages over the consolidation as a form of organization. It is easy to form, whereas there may be many technical difficulties to prevent a company from merging itself in another or even disposing of its plant and equipment The holding company also preserves the identities of the individual companies and consequently the good-will which may attach to the corporate names, and it can adapt itself to the requirements of the individual states, by incorporating a subsidiary in each state if necessary. On the other hand, it is a complex form of organization; in contrast to the concentration of responsibility and authority in the consolidation, there must be a complete administrative system for each member company. The holding company is however the most frequently used device for securing concentration of control; practically all of the largest combinations and many of the smaller ones have been built up at least in part by this means.

Mergers and consolidations. The vulnerability of the voting trust and the holding company under the anti-trust laws has had the effect of driving many who would seek relief from competitive pressure into outright mergers and consolidations. The movement toward consolidation, despite the greater technical difficulties involved, has been strengthened by the general impression that a consolidation is more likely than a holding company or a voting trust to be regarded by the Supreme Court as representing the normal expansion of the more efficient enterprises seeking larger scale operations for the economies inherent therein, without intending or effecting restriction of competition. This tendency has been reinforced by the fact that, according to the present interpretations of the Court, acquisition of physical assets is not embraced in the prohibition of the Clayton act (noted below) against the acquisition of shares of stock by one corporation in another where the effect is substantially to lessen competition.

Interlocking directorates. Another corporate device which may bring about unity of policy between ostensibly competing companies is the interlocking directorate. If the same group of men be elected to the boards of two different companies, as much uniformity in management can be secured, for the time being at least, as if one company owned the controlling stock in the other.

Trade associations. In the United States coöperation among competitors finds its most recent expression in the development of the trade association, which may be defined as a voluntary organization of independent business concerns for the purpose of furthering the interests of the particular trade or industry in which they are engaged. Though its origins may be traced back to the Civil War period, the significant development of the trade association has come since 1911. Today it is an established feature in almost every line of trade and industry. Many trade association activities, such as the conduct of research, the improvement of accounting practices, the standardization of products, the elevation of business ethics, and the development of commercial arbitration, have proved beneficial both to the industry and to the general public. However since one of the appeals to business men has been the substitution of "coöperative competition" for wasteful or cutthroat competition, it is not surprising that some of the association activities have been challenged as destructive of free competition. Thus the interchange of patent rights, credit information, and trade statistics may be made the basis of a program to eliminate price competition. Similar suspicion attaches to the so-called "open-price association," which normally provides for the filing of price lists by member concerns, the adherence to the posted price until advance notification of a change is given by filing new prices, and the interchange of full and complete information as to production, stocks, sales, and other matters necessary to show adherence to posted prices. The trade association has been extremely successful in developing an *esprit de corps*, which has much to do with the declining importance of price competition in certain sectors of American business.

"Anti-trust" legislation in the United States. American public opinion has always looked askance upon monopoly and indeed upon all large business enterprises and combinations which bear any odor

of monopolistic tendency. In the course of the rapid movement toward combination, whose outline we have sketched, there arose fear lest the entire industrial fabric become tainted with the virus of monopoly and the control of industry become concentrated in the hands of a relatively small number of men. The result was the enactment of legislation, both by the states and by the federal government, designed to break up trusts and to prevent the existence of conditions favorable to the development of monopoly. The anti-t ust laws are however a reflection quite as much of faith in competition and in the power of government to preserve it as of dislike and distrust of monopoly. In the sections which follow attention will be directed primarily to the activities of the federal government, which because of its jurisdiction over interstate commerce has been more immediately concerned in the political attacks upon the trusts.

The Sherman anti-trust act of 1890. With the exception of the Interstate Commerce Act of 1887, no legislation designed to control trusts was enacted by the federal government until 1890, when the act commonly called the Sherman anti-trust act was passed. The Sherman act is brief and exceedingly general in its phraseology. In the first section "every contract, combination in the form of a trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States or with foreign nations" is declared illegal, and according to the second section "every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a misdemeanor. . . ." Under the act the government may bring criminal prosecutions against persons guilty of violating its provisions, or it may apply to the circuit courts, by means of suits in equity, to prevent and restrain violations of the act. Upon conviction in the courts of a violation of the criminal sections of the act, a penalty, consisting of fine or imprisonment or both, may be imposed. Any person injured in his business or property by any act declared unlawful or forbidden by the Sherman act is authorized to start suit in a circuit court of the United States, and if his claim of damage be upheld in the court he can recover

three-fold the damages sustained and the costs of the suit including a reasonable attorney's fee.

Operation of the Sherman act. The criminal features of the Sherman act have not proved especially useful in preventing the growth of trusts, so difficult has it been to secure a conviction by jury on the rather vague charge of restraint of trade On the other hand in the field of forcing the dissolution of trusts, the government has met with greater, though by no means unvarying, success.

During the first ten years after its enactment the Sherman act seemed ineffective in preventing the formation of trusts or in breaking up trusts already in existence, and it was not until President Theodore Roosevelt began his energetic campaign against the trusts that there was proof that the Sherman act was still alive. In the years of his administration (1901–1909) eighteen bills in equity, twenty-five indictments, and one forfeiture proceeding were instituted Moreover new legislation was passed to facilitate the handling of anti-trust cases. To reduce the delay in settling cases an expediting act was passed which gave precedence in the federal courts to cases arising under the anti-trust act and the interstate commerce act. Another act secured the creation of a Bureau of Corporations, designed to give the government and the public authoritative information regarding the activities of the trusts.

One of the first of the significant cases to come up at this time was the Northern Securities case.[1] The Northern Securities Company had been organized as a New Jersey corporation to acquire control, by stock ownership, of two parallel and presumably competing railroads, the Northern Pacific Railway and the Great Northern Railway. The government instituted a suit in equity to have the company dissolved, and in 1904 the Supreme Court rendered its decision ordering that this be done Of still greater importance was the dissolution of the Standard Oil Company which was ordered in a decision rendered in 1911. This case marks a new interpretation of the terms of the act. According to the language of the earlier decisions all combinations or contracts in restraint of trade were held to be unlawful, but in this decision

[1] The government had already (1897) established the point that railroads were amenable to the Sherman act

the Supreme Court enunciated the so-called "rule of reason." As it is understood today, the rule of reason means that a restraint of trade which is not undue or unreasonable and which is incidental to a lawful purpose is not unlawful.

We may note one other case, that involving the Tobacco Trust. The Tobacco Trust was in part formed by consolidation, and it remained to be seen whether outright consolidation of the competing businesses would stand the legal test The court however went behind the form to the motives and results of the Trust's activities and ordered its dissolution.

The trade commission act of 1914. While considerable success had attended the efforts to dissolve trusts under President Theodore Roosevelt — and President Taft continued the program with equal energy — the feeling that more rigorous legislation was desirable was clearly reflected in the campaign pledges of the three parties actively in the field in the election of 1912. Mr. Wilson came into office pledged to "the proposition that private monopoly is indefensible and intolerable," and early in 1914 he called upon Congress to enact legislation to supplement and amend the Sherman act As a result of his recommendations two new laws, the trade commission act and the Clayton act, were passed by Congress in 1914.

Investigation of corporations. With the organization of the Federal Trade Commission the Bureau of Corporations ceased to exist, and all its pending investigations, records, and property were taken over by the commission. The act gave the Federal Trade Commission comprehensive powers of investigation and report, and granted it the authority to call upon business men to produce records of all sorts and to give testimony when required. The commission was authorized to investigate the practices of any corporation guilty or alleged to be guilty of violating the anti-trust laws, furthermore it was permitted, and in some circumstances required, to investigate the way in which a corporation was carrying out the decree of a court in an anti-trust case. But the act did not limit the commission to the investigation of infractions of the anti-trust laws, for it could call upon any corporation for information as to its business organization, methods, and relations with

other business units. This provision was inserted in order to enable the commission to accumulate a fund of knowledge with regard to the conduct of business which would be of value to Congress and to the public.

The field of business research was not confined by the law to the United States; a specific part of the duty of the commission is to keep Congress informed of trade practices in foreign countries which may influence American foreign commerce, and one of the earliest reports issued by the commission was on the subject of combinations in foreign trade with special reference to their effect on American export trade.[1] This report was incidentally a factor of considerable importance in securing the passage of the Webb-Pomerene act, which is described on page 275 below.

Advisory functions. Certain advisory functions were also provided by the trade commission act. It was felt that the Federal Trade Commission, as a result of its investigations and research, would be peculiarly fitted to assist the Department of Justice and the courts in an advisory capacity. Many of the questions which arise in connection with anti-trust cases are primarily economic, and it is reasonable to suppose that a body of economic experts may supplement the work of legal experts to advantage. With this view in mind it was in the first place made the duty of the commission, upon application of the Attorney General, to recommend a method of readjustment which would bring within the law a corporation alleged to be guilty of violating the anti-trust laws. An amicable settlement was believed in some cases to offer advantages over a long drawn out and costly law suit. In the second place, the commission was required, at the request of the court, to prepare an appropriate decree in any suit in equity brought by the Attorney General under the anti-trust laws. The court might accept the decree, modify it, or reject it entirely, but at least it would have the benefit of expert advice. Finally the commission was given the right to recommend legislation to Congress on subjects which lie in its field.

Unfair competition. It is also the duty of the Federal Trade Commission to prevent unfair methods of competition. In cases

[1] *Report on Coöperation in American Export Trade*, 2 vols., 1916.

adjudicated under the Sherman act, the courts had often been influenced in their decisions by acts which, while perhaps innocent in themselves, could not be considered lawful when practiced by a great combination in circumstances which seemed to indicate an intent to destroy competition rather than merely to gain business. Among the many methods which have been held to be unfair, we may mention: (1) price discrimination, which means that prices are cut, below the cost of production if necessary, at the competitive points while being maintained at a high level elsewhere; the Tobacco Trust and the Standard Oil Company were notorious in this regard; (2) the exclusive dealers' or factors' agreement, which compels a wholesaler to handle the goods of one concern exclusively or penalizes him, by charging him higher prices, if he carries competitive goods; (3) the use of "fighting" brands, especially designed to be sold at a price ruinous to competitors; (4) brand imitation; the American Tobacco Company thus put out at one time a "Central Union" to kill the sales of the "Union Leader" of a competitor; (5) espionage on the customers, methods, shipments, etc., of a competitor; and (6) organization of companies ostensibly independent to win the confidence of competitors.

The trade commission act created a method whereby relief from unfair competition might be obtained without the cumbersome recourse to a damage suit, which alone was possible previously. A concern which believes that it is being injured by an unfair method of competition may now lodge a complaint with the commission. If the commission finds that the method of competition in question is unfair, is practiced in interstate commerce, and involves the public interest, it serves upon the offender a complaint, giving the charges. A hearing is then held, at which time the one on whom the complaint is served has opportunity to show why a restraining order, known as a "cease and desist" order, should not be issued. If the commission still feels convinced of the justice of the complaint, it issues a restraining order to the offender, and if the order is not obeyed, the commission may apply to the Circuit Court of Appeals for enforcement of the order. The court then has complete control over the case and may confirm, modify, or set aside the order of the commission. The offender also has the right

to petition the court to set aside the order, and he has the further right of appeal to the Supreme Court.

The courts have the final decision as to whether practices condemned by the commission's cease and desist orders are actually "unfair" in the circumstances of each particular case. Thus the Supreme Court, in a majority opinion, stated that "The words 'unfair methods of competition' are not defined by the statute, and their exact meaning is in dispute. It is for the courts, not the commission, ultimately to determine as a matter of law what they include. They are clearly inapplicable to practices never heretofore regarded as opposed to good morals because characterized by deception, bad faith, fraud, or oppression, or as against public policy because of their dangerous tendency unduly to hinder competition or create monopoly. The act was certainly not intended to fetter free and fair competition as commonly understood and practiced by honorable opponents in trade." In case of non-compliance the statute provided that the commission could petition the Attorney General to begin suit in a proper federal court for a writ compelling adherence to the commission's order.

The amendment of 1938. The expression "unfair methods of competition" in the original act was interpreted by the courts as referring only to methods which were unfair to some other competitor. An amendment to the Federal Trade Commission act, passed on March 21, 1938, broadened the legal prohibition so as to include not simply "unfair methods of competition in commerce" but also "unfair and deceptive acts or practices in commerce.' Thus the commission is now able to proceed against practices inimical to consumers and the general public even though no direct injury to competitors can be shown.

This amendment also corrected an administrative weakness. Under the original act the orders of the commission were not self-enforcing, since no penalty attached to non-compliance. Hereafter the cease and desist orders of the commission become final after sixty days if no appeal to the courts is begun within that time, and after that period penalties for non-compliance apply automatically.

In the past much of the commission's work has been directed against false advertising as an unfair method of competition. The

amendment of 1938 enlarged the powers of the commission in this field by prohibiting the dissemination of any false advertisement to induce the purchase of food, drugs, devices, or cosmetics. An advertisement is "false" if it is misleading in any material respect or if it fails to reveal material facts.

The Clayton act (1914) and the Robinson-Patman act (1936). The Clayton act of 1914 prohibits, with certain limitations, local price discrimination, tying contracts, holding companies, and interlocking directorates, the intention being, by making these practices specifically unlawful, to stifle the creation of trusts, conspiracies, and monopolies in their incipiency and before their consummation.

It is made unlawful for a person engaged in commerce to discriminate in price between different purchasers of a commodity, where the effect of such discrimination may be substantially to lessen competition or may tend to create a monopoly in any line of commerce; provided that discriminations may be made for differences in the grade, quality, or quantity sold, or in the cost of selling or transportation, or when made in good faith to meet competition, and that a seller may select his customers in bona fide transactions and not in restraint of trade.

In 1936 the prohibition against price discrimination was elaborated by the passage of the Robinson-Patman act, which provides that it shall be illegal either to grant or to receive, directly or indirectly, any discrimination in price or service in any sale of commodities in interstate commerce. Price differentials may make due allowance only for differences in the costs of manufacture or sale to different purchasers. Furthermore the Federal Trade Commission may fix and establish quantity limits if price differentials on purchases in larger quantities would be discriminatory or promotive of monopoly. Whenever price discrimination is found, the commission is to order its removal, and the burden of proof in defense of discrimination in prices, services, or facilities rests upon the person charged with the discrimination. The difficulties and complexities inherent in the application of these statutory provisions have raised serious doubts as to the possibility of effective administration.

Tying contracts, which bind the purchaser or lessee of certain products to buy or use certain other products exclusively, are declared illegal by the Clayton act, where such lease, sale, or contract for sale may tend substantially to lessen competition or to create a monopoly in any line of commerce.

No corporation engaged in commerce may acquire, directly or indirectly, the whole or any part of the stock of another corporation also engaged in commerce, where the effect of such acquisition may be substantially to lessen competition between the two corporations, or to restrain such commerce in any section of the country, or to tend to create a monopoly in any line of commerce. Holdings of securities already acquired are not made illegal, and a corporation may continue to acquire securities for investment or for other purposes, provided it does not offend in the above particulars.

No person may at the same time be a director in two or more corporations engaged in commerce (other than banks and common carriers), any one of which has capital, surplus, and undivided profits exceeding a million dollars, if such corporations are or have theretofore been competitors so that the elimination of competition by agreement between them would constitute a violation of any of the provisions of any of the anti-trust laws.

The provisions of the Clayton act are to be enforced by the Interstate Commerce Commission when they apply to common carriers, the Board of Governors of the Federal Reserve System when they apply to banks and trust companies, and the Federal Trade Commission when they apply to other corporations. The manner of procedure in each case is specifically outlined.

Remedies provided by the anti-trust acts. In addition to proceedings before the administrative boards under the amendments of 1914, the statutes have provided other methods of proceeding against violations of the anti-trust laws. Criminal prosecutions may be started by the government, charging monopoly, attempt to monopolize, or conspiracy in restraint of trade. The government may initiate suits in equity, seeking the dissolution of combinations existing in violation of the law, or injunctions against specific acts or agreements in restraint of trade. Or the government may de-

mand the forfeiture of goods being carried in violation of the law when moving in interstate commerce. Individuals and corporations may sue for recovery of three-fold damages for injuries resulting from violations of the law, and under the Clayton act individuals and corporations are permitted to seek injunctions against losses threatened through violation of the anti-trust laws.[1] The Clayton act provides that if, in a suit brought by the United States, it is decided that the defendant has violated the anti-trust laws, this fact must be accepted as prima facie evidence in any suit brought by another party under these laws, thus making it easier and cheaper for the injured person of small means to sue a trust.

The Webb-Pomerene act. In 1918 the trust policy of the United States was modified by the passage of the Webb-Pomerene act, which exempted from the provisions of the anti-trust acts combinations engaged solely in export trade, provided that such combinations do not restrain trade within the United States nor employ unfair methods of competition with American competitors. This liberalization of the anti-trust laws was conceived to be necessary to enable American exporters to compete with those of countries where combinations are either tolerated or encouraged. The result has been to allow small manufacturers, who otherwise might find it difficult to break into the foreign markets, to enter as members of combinations. The larger manufacturers have shown little inclination to form export combinations, and it is doubtful whether this act will ever succeed in stimulating export trade to the degree anticipated by its sponsors.

Achievements of the United States anti-trust program. Though little effort at enforcement was made in the years immediately following its enactment, the Sherman act and its amendments have been a continuing threat to monopolistic practices since the turn of the century. The government has won a number of spectacular victories in the courts, securing the dissolution of some of the most conspicuous trusts and combinations. In many industries of national importance the threat of anti-trust proceedings has

[1] It may also be noted that violation of the statutes may be cited as a defense in a private suit brought to enforce a contract alleged to be in violation of the law, and criminal proceedings may follow a violation of an injunction issued by a court in an anti-trust case.

sufficed to secure voluntary agreements, or "consent decrees," whereby the alleged offenders have accepted court orders binding them to desist from specified practices criticized by the government. The large and powerful business units, being more exposed to the critical appraisal of the public, have found themselves compelled to forego the ruthless tactics of the early Standard Oil and American Tobacco trusts, and policies of coöperation, "live-and-let-live," and persuasion have become the usual methods of rendering competition less distasteful to industry. The Federal Trade Commission has been successful in eliminating many unfair competitive practices and raising the level of business ethics.

Yet such results as have been achieved have scarcely retarded the trend toward combination and concentration of control in American business, and there is much evidence pointing towards a marked decline in price competition in many lines of industry. With a reduction in the number of competitors, conditions of imperfect competition become more pronounced, and non-price competition tends to replace price competition, with emphasis on product differentiation, sharing the market, price leadership, price stabilization, and increasing sales pressure. The increasing influence of trade associations, aided by the philosophy of the NRA, has developed sentiment in favor of coöperation in place of competition and restriction of output to avoid glutting the market and undermining the price structure.

In view of the avowed objective of the federal legislation to prevent monopoly and preserve competitive conditions, it must be admitted that the regulatory efforts have at best attained only partial success, since the problem is still with us and in some respects in aggravated form. Nevertheless comparison of existing conditions with those which would presumably have prevailed in the absence of governmental interference and regulation is sufficient to indicate a real accomplishment to the credit of the anti-trust program.

Reasons for moderate results of the anti-trust program: Is monopoly inevitable? The many forces which appear to be tending toward monopoly and the moderate degree of success achieved by the government in opposing such forces by restrictive legislation

raises the interesting and far-reaching question whether there is after all a fundamental tendency toward monopoly which it is idle to resist — whether in short monopoly is inevitable. If it be true that economic forces, operating in a social system based as ours is upon personal liberty, private property, and individual control of enterprise, tend inevitably to bring about a condition of monopoly in the production of a substantial part of the basic products which mankind requires, then we are faced with a human problem of the first magnitude, a problem which may well put upon the defensive the very fundamentals of the existing social order.

Forces favoring monopoly. At first glance the forces favoring monopoly appear almost irresistible. There is first of all the well known economy of large-scale production in many lines of industry. There are the economies which in certain cases appear to accompany the combination of separate business units under a single management. There is the occasional appearance of cutthroat competition, an unstable situation leading almost inevitably to monopoly. There is the constant effort of the captains of industry, when competition becomes especially keen and irksome, to find escape through some form of combination, making use of every possible device — pool, trust, holding company, and so on — which human ingenuity with millions at stake can devise. The corporate form of organization is a helpful instrument. The protective tariff, while probably not the "mother of trusts" as it has sometimes been called, has unquestionably reduced the competition which American producers have had to face and in that respect has cooperated with the other forces tending toward monopoly. And besides the desire to secure monopoly profits there is also the urge to secure promoters' profits by the organization of large combinations which may have monopoly power.

Acquaintance with economic history will convince one that the enactments of legislative bodies may be futile if they are in opposition to powerful economic forces. If cutthroat competition is an inevitable feature of large-scale production, it would be the part of wisdom for society to seek stabilization in monopoly and to direct its efforts to regulation calculated to enforce good service and reasonable prices. If a combination of producing units under one

management is more efficient than when each of the units is under separate management, if a combination of combinations in exclusive control of the field is the most efficient of all, it is folly to try to run counter to the natural tendency of business men to seek the most profitable form of organization. Society cannot permanently reject the most economical type of productive organization. This is virtually the conclusion which we have reached as regards the railroads, and our railroad policy and legislation are based largely on the assumption of the inevitability of monopoly.

Limits to monopoly. But the conditions which prevail in the realm of transportation, communication, and the municipal public utilities, which confessedly point the way to ultimate monopoly, are not universal throughout the whole field of production.

At the forefront stands agriculture, most basic of all basic industries, where economy is not to be found in large-scale production or combination, and where fundamental economic laws seem destined to preserve indefinitely the *régime* of small productive units and free competition.

In the fields of manufacturing and marketing it is doubtful if the monopolistic combination is in many cases the most efficient form of organization. We have already learned that the efficiency of large-scale production and combination has very real limits. There are few important lines of industry in which this limit would not be reached long before the would-be monopolist had become great enough to absorb the whole. Even assuming the monopoly to become established, lethargy in management and lack of progress in the technique of production and the methods of organization will tend to appear, so that the monopoly which at the moment of its formation is representative of the most efficient methods of production carries with it no assurance of continued efficiency.

We conclude that there is little reason, from the standpoint either of theory or of history, for believing that the economies of combination alone are sufficient to warrant the growth of monopolies in many fields of production. Nor is there evidence to show that competition between large units must necessarily be of the cut-throat variety. This is essentially the result of an unbalanced condition in production, where the total productive capacity in a

given industry is greater than the market requires. This situation is more or less normal in certain lines, such as the railroad business and the public service industries in general; here monopoly may frequently be inevitable and, under adequate regulation, in the public interest. As to industry in general, apart from occasional excess capacity resulting from mistaken judgment or depression conditions, there is little reason to anticipate that such a condition will present a constant menace to industry. It is more characteristic of the general situation that, as each period of intense competition weeds out the weaker producers, industry settles down to a normal condition of competition between a few strong rivals, no one of which has reason to hope for monopoly control.

Standards set by legislation. The present legislation, as it is interpreted by the courts, places great stress upon the question of intent on the part of the corporations affected. Thus when a case comes before the courts it is not enough to consider merely the results of the combination; the intent to achieve an illegal combination or restraint must be shown. The demonstrable economic facts — whether the combination realizes important economies and whether it creates opportunities for monopolistic restraints — the only objective criteria by which the law can be administered, are not conclusive in proving a violation of the law. The absence of objective criteria fosters protracted and expensive litigation and imposes upon the Department of Justice the impossible task of ferreting out evidence as to the unexpressed intent of those who formulate and control corporate policies. The cost of preparing a single case may be from \$100,000 to \$150,000 or even more and may require the full time of a dozen attorneys for months or years.

The task of administration. An important cause of the meagre accomplishment of the federal anti-trust legislation is to be found in the very magnitude of the task in relation to the administrative machinery that Congress has provided. Initiative in administration centres in the Anti-Trust Division of the Department of Justice, which is responsible for suits arising under some thirty-one important federal statutes. With its present staff and multitudinous responsibilities, the division finds it impossible to prosecute,

or even to investigate, all of the alleged violations of the anti-trust laws that are reported to it.

Influence of patent laws. The patent laws have in practice sometimes served the ends of the monopoly builders. When the owner of a patent licenses others to manufacture under his patent, he may specify the quantity of the patented article to be produced, the markets in which the article may be sold, and the price to be charged. The patent laws have sometimes stifled new competition at its inception, by limiting the market for patented improvements on existing processes or mechanisms to the owners of the basic patent. Newly patented articles which might provide competition with products in current use have on occasion been bought and suppressed by corporations financially interested in established products or technical processes. Thus the patent laws, whose basic purpose is frankly to grant certain monopoly privileges, have in some cases promoted the development of monopolistic industry beyond what the spirit of patent legislation contemplated.

Contradictory laws. A loss of faith in the efficacy of competition is reflected in the enactment of laws by the federal and state governments in conflict with the basic assumptions of the anti-trust legislation. While the national industrial recovery act provided that monopolistic practices should not be sanctioned, many of the codes were approved with provisions for resale price maintenance, restriction of output or capacity, open-price plans, and other devices restrictive of competition. The stimulus to the formation of trade associations and the extension of their activities has lasted beyond the demise of the NRA and has been productive rather of coöperation than of competition.

At the present time serious threat to the continuance of competitive conditions is to be found in a recent epidemic of state "fair trade laws." These laws permit manufacturers of branded or trade-marked articles which are in "fair and open competition with commodities of the same general class" to enter into contracts with dealers for the maintenance of designated prices in the resale of such articles. Similarly in conflict with the spirit and the prevailing interpretation of the federal anti-trust statutes is the Miller-Tydings act, passed by Congress on August 17, 1937, which

exempts such price maintenance contracts from the prohibitions of the anti-trust laws.

Proposals for reform. There are those who interpret the record of the United States government in administering the anti-trust legislation as indicating that government regulation of private industry in the interest of preserving competitive conditions has had a fair trial in America and proved itself a failure. A more reasonable interpretation, it would appear, would take into account the limitations in the laws, the handicaps which have beset their administration, and other adverse circumstances and would conclude that the time has not yet come to abandon the hope that government may yet be able to devise measures of control which will prevent the formation of monopolies and preserve competitive conditions generally in American industry and trade.

As a first prerequisite of successful attack upon this problem, we need more knowledge. We do not at present know enough about the working of our trade regulations, the nature and functioning of competition in specific industries, or the organization and practices of particular trades. This knowledge is essential to the formulation of a sound and comprehensive program of reform. Presumably the investigation started in the summer of 1938 will furnish information on these points, but provision should be made to continue this type of research. This might well be expanded to provide regular collection and dissemination of trade and industrial statistics for the information both of industry and of consumers. If this were done, a function frequently associated with efforts restricting competition among members might be removed from the trade associations.

If future efforts to control industrial monopoly are to be successful, they must in future at least avoid those past errors which experience has disclosed. Objective standards of business conduct, looking to economic consequences rather than "intentions," should be clearly set up in the laws. Contradictory laws, which interfere with the main objective, should be scrutinized and, when not required on other grounds, should be repealed. The direct and indirect effects of existing patent laws need fresh examination.

The enactment of a federal incorporation law, making a federal

charter a requirement for all business subject to federal jurisdiction and permissive for others, is a suggestion that merits consideration. The significance of such a proposal can be appreciated when it is realized that many of the corporate privileges now conferred by the states, such as the right to own shares of stock in other companies, to engage in a bewildering array of undertakings, or to issue a wide variety of non-voting securities, are the tools of corporate empire building and that if any of these powers need to be withdrawn or limited, such action would be facilitated by uniform incorporation laws designed and administered by the national government. On the other hand, there are serious considerations that may be urged in opposition to federal incorporation.

Conclusions. To assert that the survival of the system of private capitalism – perhaps even of political democracy itself – requires the control of monopoly and such regulation of trade and industry as will maintain competitive conditions in a reasonably large portion of industry is not an overstatement of the case. The efficient regulation of trade and industry demands legislation soundly conceived and drafted, administration by competent and honest public servants, and the support of an alert and intelligent public opinion.

The problem is not a simple one. Conditions vary from industry to industry and an effective program of regulation must have sufficient flexibility to adjust itself to many situations. There is even reason to expect that experience may find instances, perhaps not now suspected, of industries to which competitive controls are not appropriate. Here the extension of regulation such as is now applied to the public utilities may be the order of the future; or it is possible that the use of government corporations, with representation for investors, workers, and the general public, may prove more efficient than government regulation and less costly than government ownership. Past experience also suggests that laborers, consumers, and investors may require additional protection from some of the undesirable results of competition. But despite exceptions and difficulties, there is reasonable hope for the success of a revised and vigorous program. Certainly the importance of the goal warrants the attempt.

XLII

MARKETING OF CORPORATE SECURITIES AND OTHER PROBLEMS OF THE CORPORATION

Marketing corporate securities: Middleman services. The opportunity afforded investment bankers and others to secure profits from the issuing of corporate securities has been suggested as the third motive leading to the formation of large combinations. Many corporations would encounter serious difficulties in securing the funds which they need for expansion of their operations if they did not have the assistance of a multitude of specialized banking houses, investment brokers, and other middlemen standing between themselves and the investing public. The services of middlemen are not always required. Many a corporation has a very small capitalization and does not try to interest the public in its stock; the small group of men who have organized the corporation and who expect to become its responsible officers subscribe for the whole issue. Other companies are local in character and are able to appeal to local investors without the assistance of the middleman Even in some large concerns the ownership of the stock is concentrated in a few hands; the stock of the Ford Motor Company is held entirely by the members of the Ford family. Yet in most large corporations there is a wide distribution of ownership – there were for example 235,000 owners of stock in the United States Steel Corporation in 1935—and a company with a capitalization of several millions is frequently too large to be financed either by a group of individuals or by a small community. An appeal must be made to a wider public, and bond houses and investment bankers, brokerage houses and the stock exchanges exist to carry that appeal to investors.

Selling stocks and bonds is a type of merchandising, a highly specialized type, in which a detailed knowledge of the goods to be sold and a wide acquaintance among buyers are prerequisites for success. The corporation has neither the organization nor a suffi-

cient knowledge of the market for selling its securities. True, it could install a security sales department, as some companies have done for the sale of stock to their employees or to their clients, but the occasions when it has securities to sell are rare; *i.e.*, when the corporation is formed or when new issues are floated to secure additional capital. Also it would ordinarily take the corporation a considerable time to sell its securities, and since it presumably needs the funds at once, this may decide it to put the matter in the hands of bankers who will agree to hand over a certain sum of money by a certain date. Furthermore, if the corporation experiences any difficulty in disposing of its securities, it may acquire a bad name which will militate against success in marketing future issues. The financial middlemen are constantly dealing in a wide range of securities; they are in contact with investors all over the country and can handle the whole business more cheaply and expeditiously than the corporation itself.

Underwriting bonds. It is not properly a function of an elementary text in economics to discuss in detail the types of investment houses and the special functions of each type. We must treat the subject in a very broad way, and we shall use the terms "investment houses" and "investment bankers" loosely to include all the houses which assist in selling securities to the public.

Let us assume that a corporation has submitted a report to a prominent banking house which specializes in financing investments in industry and has satisfied it that the purpose for which the proceeds of a proposed bond sale are needed is sound and likely to be profitable. There are several million dollars of bonds to be sold, and the banking house is unwilling to assume all the responsibility for their sale. It therefore organizes a syndicate or group of investment bankers, located in the same city or in other cities, which agrees to underwrite the sale of the bonds.

"To underwrite is to insure, and as the term is used in the financial world, *underwriting may be defined as insuring the sale of a corporation's securities in advance of a public offering.*"[1] Practically the syndicate agrees to take the issue of bonds and make payment for it at a specified time and at a fixed price per bond.

[1] L. H. Haney, *Business Organization and Combination*, 1913, p. 301.

Naturally it expects to sell the bonds before the date when payment is due, but if it fails to do so or if it fails to sell the bonds at the price anticipated, the loss falls on the syndicate and not on the corporation. The assumption of this risk is the essential feature in the underwriting agreement.

The syndicate expects in the natural course of events to make a profit from handling the bond issue, and if it agrees to pay the corporation ninety-five dollars per bond, it may fix upon ninety-eight dollars as its selling price to the public, thus leaving it a commission of, let us assume, one dollar for selling and two dollars for underwriting the sale of each bond.

Each member of the syndicate agrees to be responsible for the sale of a given portion of the total bond issue and therefore has an incentive to bring into play all of its own sales organization and to interest the smaller houses with which it has dealings. Advertising and personal solicitation by salesmen play their part here just as they do in selling any kind of merchandise. Naturally the wider the appeal which is made to the public, the greater is the chance to dispose of the bonds before the date when payment has to be made to the issuing corporation. The financial organization is so closely tied together that for the sale of a very large issue of bonds it is possible to interest bond houses and investment bankers in almost every city of the United States.

The investment houses act as middlemen; they do not expect to tie up their own capital in these securities, but sometimes that is unavoidable. If the issue has not been sold before the day arrives when the corporation must be paid, the syndicate must make payment in full, either from its own resources or more commonly by borrowing, pledging the unsold bonds as collateral with the banks. Meanwhile the sale of bonds goes on, and as funds are received they are used to reduce the amount of the loan.

Marketing an issue of stock. An issue of stock may be sold by the use of this same machinery; that is, by a public offer and subscription. Frequently however common stock is acquired by investment banking houses which are active in the promotion of a new corporation and is then distributed to the investing or speculating public by way of the stock exchanges. The extent

to which this method is used is suggested by the report that in 1928 the firm of J. P. Morgan and Company marketed, by public offer, sixty-eight issues of securities of which not one was a common stock; in the same year the eight Wall Street houses which led in the floating of securities are said to have distributed by public offer 995 issues, of which only 24 were of common stocks.[1]

The stock to be resold may be acquired by purchase, as in the case of the financing of the Alleghany Corporation, where the Morgan firm bought 1,250,000 shares at twenty dollars a share. It may, on the other hand, be acquired in exchange for property of some sort, such as stock in another company, or in return for services rendered.

Promoters' profits. Profits which accrue to investment bankers from underwriting an issue of corporate securities are usually quite modest in comparison to the return which may come to the promoters of a new corporation or of a combination. The organization of a large corporation or of a combination involves a multitude of details, such as the securing of a charter, the determination of the financial structure of the new company, and perhaps the underwriting of issues of securities. In the organization of a combination it is also necessary to negotiate for the purchase of a controlling interest in the securities of those companies which are to be brought under common management. Frequently this control can be acquired by an exchange of the securities of the holding company for those of the constituent companies, but in some cases the stock must be purchased for cash either on the open market or from holders of large blocks of stock.

For these services the promoters — and prominent among them are the large investment houses — may be paid in cash, but it is a common practice to reimburse them in whole or in part in common stock of the new corporation. Where payment is made in common stock it is of course true that the present worth of the payment is contingent on the expected profitableness of the newly organized corporation. In future years the stock may be worth little or nothing, or it may be worth a great deal.

[1] J. T. Flynn, *Security Speculation*, New York, 1934, p. 140.

This has an appearance of fairness, for promoters must have faith in their venture to risk their services for a reward whose market value may not be determined for some time to come. Undoubtedly in many instances this is true. On the other hand, we should not be blind to the fact that a return may be secured which bears no ascertainable relationship either to the worth of the services performed or to the risk involved. Perhaps the most familiar example is afforded by the United States Steel Corporation. The reader will recall that the formation of this company involved the acquisition of a controlling interest in the stock of a number of independent companies. This was acquired by an exchange of the bonds and stock of the new corporation for stock of the member companies. In addition the underwriting syndicate in charge incurred expenses of $28,000,000, of which $25,000,000 went to the new corporation for working capital and $3,000,000 were used for organizational expenses. "In return for this outlay and its underwriting services the syndicate was given by the Corporation a total of 1,299,975 shares of stock (half preferred and half common), having a par value of $129,997,500. On this stock the syndicate realized about $90,500,000. Deducting the cash expenditures ($28,000,000), it is evident that the syndicate made a profit of $62,500,000. Of this amount one-fifth went to the firm of J. P. Morgan and Company, the syndicate managers, for their services; and the balance ($50,000,000) was distributed among the members of the underwriting syndicate, including the Morgan firm." [1]

In the current discussion of the profit or loss to investors in the common stocks of such combinations and in the debate as to whether promoters' profits are or are not exorbitant payments for the services they render, one is apt to lose sight of the more important fact that the possibility of securing such large profits from the sale of securities leads bankers to a restless search for opportunities to bring about still other corporate reorganizations.

[1] E. Jones, *The Trust Problem in the United States*, New York, 1921, p 288. An interesting and reliable account of promoters' profits in the period prior to 1920 will be found in Chapter XII of this study. For accounts of more recent financing, such as the financing of the Alleghany Corporation, Insull Utility Investments, Inc., etc., the reader is referred to J T Flynn, *Security Speculation*, 1934, pp. 141 ff.

In other words, monopolistic aggregations of capital may be brought into being by the operation of a motive quite remote from the desire to secure monopoly profits, though the possibility of monopoly profits in the future may be one of the reasons leading to a favorable judgment of the combination.

Promoters' profits, in so far as they arise from a sale of stock, depend upon a buoyant security market, one characterized by the most optimistic discounting of the future earnings of the corporation. The promoters may be fortunate in finding that speculative fever has worked things up to the point where almost any security will find eager purchasers, but even if this be not true, many methods have been available to "create a market" for the stock. Pool operations to help establish a price for the stock, optimistic information, sometimes even misleading, disseminated as to the future prospects of the company, and other methods have been used in the past to induce speculators and investors to take the stock at prices profitable to the promoters. Measures which tend to curb speculation and to check certain undesirable practices on the stock exchange should assist somewhat in preventing the formation of combinations by checking the manipulation of security prices to artificially high levels and thereby reducing the promoters' profits from the sale of stock.

The services of investment bankers. Directly or indirectly a great many different individuals are adding each year to the total amount of capital which producers and others are using. The investment bankers act as middlemen by directing the flow of this capital to one enterprise or another. In this they do not play a passive rôle, for the risks inherent in underwriting require them to scrutinize carefully each proposition before committing themselves to it, and a nice regard for their professional repute should compel such precaution. Their wider information regarding investment possibilities and their more profound knowledge of finance tend to prevent misdirection and dissipation of capital. By refusing to underwrite securities for corporations in a line which is already overdeveloped or by refusing to handle the issue of a corporation in whose officers they have no confidence, they can direct the flow of capital and throw the control of industry into

the hands of men best equipped to manage it. Likewise they can and should reject the plea of a government to float an issue of bonds when investigation shows the finances of the government to be in doubtful condition.

Investment bankers are consequently in a position of great public responsibility which unhappily they have not always adequately filled. The speculative madness which in past eras of prosperity has seized upon the multitude did not find them immune. Hundreds of millions of dollars in stocks and bonds, domestic and foreign, have been sold to the investing and speculative public without sufficient investigation on the part of the underwriters into the affairs of the corporation or of the borrowing government. In many such cases there was eventual loss of all or a large part of the investment. Houses competed with each other for the privilege of floating loans for South American governments, German municipalities, and private corporations, in order to secure the profits which came from the sale of these securities to a public whose appetite was at the moment without apparent limit.

The investment bankers and the allied agencies are however an essential part of our economic system. They have been called into being to assume burdens which were becoming too heavy for the corporation, and they can perform the essential function more cheaply than can the corporation. Their function is one of immense importance to society as a whole; it can be performed by them in a satisfactory manner, and it is incumbent upon society to demand that this be done and to take such measures as seem necessary to assure it.

Legal regulation. Securities act of 1933. The belief that conditions surrounding the issue of securities required control led to the securities act, passed by the United States Congress in 1933 and revised by amendment in 1934. Save for railroad securities, which remain under the control of the Interstate Commerce Commission, this legislation regulates the conditions of public offering of all securities with a maturity of more than nine months issued by any foreign agency or domestic business corporation. The first steps in the public sale of any such security are the regis-

tration of the issue with the Securities and Exchange Commission (SEC), created in 1934, and the filing with the commission of a statement of all facts relevant to the determination of the investment value of the security. The purpose of this is to give the investor certain knowledge which was seldom offered him in the past and to safeguard him from incomplete or vague statements regarding the issue. Direct protection against misleading statements is found in the provision that, after a statement of operations for twelve months has been issued, a purchaser of a security may collect damages if he can prove that he relied in his purchase on an untrue statement in the registration papers or the attendant literature. Underwriters, officers, or directors are not however liable for errors in a statement due to mistakes of experts if they can prove they had no reason to believe that the items in question were untrue. The responsibility of underwriters is limited to the parts of an issue they respectively handle, and a plaintiff seeking to recover damages under the act may be required by the court to post a bond to cover the costs of the suit. These provisions afford a very material safeguard to the investor, who must in the very nature of the case rely on the statements of investment bankers in making his choice of investments.

The stock exchange: Description. No discussion of the marketing of securities could be considered adequate without mention of the facilities of the stock exchanges for the quick disposal of stocks and bonds of the larger corporations and governments.

On June 30, 1937, there were twenty-two exchanges registered with the Securities and Exchange Commission as "national securities exchanges" besides seven others exempted from registration. For purposes of illustration we may take the New York Stock Exchange, which handles over 65 per cent of the share transactions concluded on all stock exchanges in the country. The New York Stock Exchange dates back to 1792, when a few brokers made an agreement regarding their relations with their customers. From this small beginning, with transactions confined to a few government issues and bank and insurance stock, there has been a gradual development in the scope of operations and in the volume of business transacted. This exchange is a voluntary association, with an

authorized membership in 1938 of 1,375. It conducts business according to rules and regulations adopted partly in conformity to legislation and partly to preserve a high code of business ethics. A governing committee is clothed with power to see that this code is adhered to, and for serious infractions a member may be expelled and his "seat" vacated. No one not a member of the Stock Exchange is allowed to do business on the floor of the exchange. Most of the members are brokers and buy and sell at the order of a principal, although there are some who trade independently, buying and selling on their own account.

A stock or bond must first be listed on the exchange before it can be dealt in on the floor, and as a preliminary to listing it must have been sold by the corporation concerned. The exchange does not float new issues of stocks and bonds. Listing further involves an application by the officers of the corporation, which is referred to a committee of investigation, according to whose report the application is either approved or denied. A stock which is listed on the Stock Exchange may seem to the public to bear a certain stamp of approval, but actually the Stock Exchange is presumed to do no more than credit the management with honest intentions in the sale and disposition of its stock. It never guarantees the worth of the stock or bond so listed, nor does it recommend its purchase. The Illinois Securities Law of 1919 quite aptly summed up the significance to be attached to listing, when it classified securities which are dealt in on the New York and certain other exchanges in that group of securities whose inherent qualities ensure their sale and disposition without fraud.

Services of the stock exchange. A stock exchange can be expected to perform three important functions. In the first place, it provides a ready market for the purchase and sale of listed securities. Such a market, where securities can always be sold at some price, tends to make investment in listed securities more attractive. Also, since bankers know that they can always sell listed securities if necessary to protect themselves, investors can use their securities as collateral for bank loans if they do not choose to sell them. In the second place, the exchange furnishes a market where, by a continuous process of buying and selling, prices of securities are determined,

representing their true values based on the capitalization of future income-bearing possibilities. And finally, as a result of this continuous evaluation of securities, the stock exchange assists in directing the flow of capital into those industries where it can most profitably be used.

The importance of these services and the prominent place which the financial news gives to activities on the organized exchanges, particularly the New York exchange, should not however blind us to the fact that a vast amount of investment takes place outside the exchanges. A great volume of funds is invested in real estate and in real estate mortgages without the medium of an organized exchange. There are hundreds of thousands of corporations which have no securities listed on any exchange; these corporations in fact control more wealth than the thousand or so which are listed on the New York Stock Exchange. Furthermore few of the great corporations now represented by listing on the exchanges seem to have acquired any substantial portion of the capital necessary for their development through the mechanism of the exchanges. Much of the stock issued in the promotion of mergers and reorganizations has indeed come into the hands of the public through the exchanges, but it is a question as to how far this stock has represented the investment of new capital as distinguished from promoters' bonus.

Criticisms directed against the exchanges. The economic problems arising from the securities exchanges do not come from the fact that they are an essential prerequisite to investment on a large scale, but rather from the fact that an exchange which fails properly to perform its functions is a disrupting and disturbing factor in economic life. We could conceivably get along without stock exchanges, but we cannot run smoothly with malfunctioning exchanges.

In the United States the stock exchanges have for some time been a cause of controversy, and in the years following the crisis of 1929 criticism of the securities markets in general and of the New York Stock Exchange in particular was intensified. The opinion was widely held that the New York Stock Exchange was conducted in a manner more suited to a private club than an organization

performing essential economic functions. More specifically it was charged that the operations of the exchanges were inimical to economic and social progress because they tended to promote excessive speculation and to divert excessive amounts of credit from trade and industry to the securities markets. Low margin requirements and the ease with which credit was available to operators in the securities markets were blamed for much of the allegedly excessive speculation. The operations of pools, leading to sudden and unpredictable fluctuations in the prices of securities, were cited as evidence of the use of the machinery of the exchanges by gamblers intent on milking the public. The practice of short selling, accompanied by bear raids, was said to drive securities prices to artificially low levels with great loss to investors forced to sell their holdings. The specialist, whose "book" contains information as to the buying and selling orders awaiting execution when the market price of the security shall reach designated points, was criticized for using his confidential information to guide his own trading activities or for making such information available to pool operators. And finally criticism was directed at the publicity policies of the exchanges, especially the failure to make available significant facts concerning securities listed on the exchanges. This criticism resulted finally in legislation by Congress in 1934.

Legal regulation: Securities exchange act of 1934. This act institutes federal control of security exchanges and over-the-counter markets operating in interstate and foreign commerce and through the mails. It is administered by the Securities and Exchange Commission, of five members appointed by the President with the consent of the Senate. For purposes of administration of the securities act and the securities exchange act the commission has divided the country into eight zones, with a regional office in each; in addition it maintains a field office in Washington, with jurisdiction over the exchanges in the District of Columbia and certain of the neighboring states. Exchanges must be licensed for operation. Corporations with securities listed on an exchange must file complete organization and financial statements with the commission and with the exchange, and the former may require an annual audit and the submission of quarterly financial statements.

Officers, directors, and owners of more than ten per cent of the stock of a listed company must provide the commission with statements of their holdings and make monthly reports of any changes. An attempt is made to render unprofitable the use of "inside information" by such persons in the purchase or sale of stock of the corporation through the provision that the profit realized within six months of a purchase or sale of such stock is, subject to some limitations, recoverable by the corporation.

While the commission is granted sweeping powers to prohibit any manipulative or deceptive practice, certain practices are definitely outlawed, and others are suggested by name as needing regulation or prohibition. Force is given to the prohibitions of the act and to the regulations of the commission by providing that a violation of these prohibitions or regulations of manipulative practices make the violator liable to suit for the recovery of damages which the purchaser or seller may suffer as a result of such violation. The manipulative practices which are unconditionally forbidden are: fictitious transactions or "wash sales"; transactions to raise or depress prices to induce the purchase or sale by others; circulation by the seller or purchaser of information as to a possible change in price resulting from market manipulation or activity; the making of statements to influence a purchase or sale when there is reasonable ground for believing such statements to be false or misleading with respect to any material fact. The commission is empowered to prohibit or regulate certain other practices.

Power to regulate credit advanced for margin accounts is given the Board of Governors of the Federal Reserve System, and brokers must borrow from member banks of the Federal Reserve or from such others as comply with the regulations of the Board of Governors. The purpose of the first provision is to require a more substantial deposit on the purchase of a security than was common during recent boom times and thus to eliminate the small speculator on margin and to reduce the total volume of margin accounts. The second provision is intended to give a greater degree of control over the expansion of credit for speculative purposes by curbing the practice of banks of borrowing from large corporations having idle cash reserves in order to relend the money to speculators on margin

accounts. In addition the commission has power, subject to certain restrictions, to limit the aggregate indebtedness of any broker. The anticipation is that this regulation of credit will promote greater stability of prices on the exchanges and eliminate "runaway" markets. Finally we should mention the power granted the commission to segregate and limit the functions of members, brokers, and dealers. Important among the objects of such regulation are limitation of floor trading by members, and the definition and segregation of the functions of certain types of floor traders.

The SEC and the exchanges. The activities of the Securities and Exchange Commission in the regulation of the organized securities markets have been too specialized to justify detailed treatment here. The first step in the regulatory process, the registration of the exchanges, has been effected. Since the continuance of registration is conditioned upon the filing of supplementary reports, the main objective of the act of 1934 -- to make available reliable, current, and comprehensive information regarding the affairs of any company with securities listed on the exchanges -- has been accomplished.

The rules and regulations governing activities on the exchanges have received the careful attention of the commission with reference to the performance by the exchanges of their essential services to the financial and investing public. The guiding philosophy of the SEC thus far appears to be to secure the largest possible degree of self-regulation by the exchanges themselves, recognizing that such control can be swift, flexible, intimate, and preventive, whereas detailed supervision by the commission might be more time-consuming, costly, and productive of friction. The result of this policy is illustrated by the reorganization of the New York Stock Exchange, following a warning that increased governmental regulation could be avoided only if the exchange were so reorganized as to permit effective regulation of its own members and more adequate response to the needs of the investing public. The Board of Governors of the Exchange must hereafter include representatives from all classes of exchange members as well as three representatives of the general public. The president of the Exchange is now a full time salaried executive, who must divest himself of all other

business interests and devote full time to the administrative problems of the Exchange; he appoints all important committees and is himself an *ex officio* member of all committees. The example of the New York Stock Exchange has been followed by the other leading securities markets.

The formal reorganization of the exchanges is preliminary to further changes in the organization and operation of these bodies. Proposals for broadening the membership while reducing the number of members, for floor policing, for the safekeeping of securities and customers' funds, for regulating short-selling and the activities of specialists, for the incorporation of member firms, and for requiring larger capital for such corporations have received consideration and in some instances been put into effect.

Conclusion. Undoubtedly this legislation will result in a reduction of the volume of trading on the exchanges, but it is anticipated that the trading thus lost will be of the speculative type which has profited and lost by the creation of artificial prices. A market supported chiefly by investors rather than by speculators might be more sluggish, but it should serve all legitimate needs and in addition give the normal forces of demand and supply an opportunity to determine prices more truly representative of future earning capacity and therefore more reliable as a director of the flow of capital. Reduction in speculative profits might also diminish the effectiveness of one of the incentives towards the formation of monopolistic combinations. It will however be a matter of years before any mature judgment can be reached of the appropriateness of the securities act and the securities exchange act to achieve their purposes; perhaps in the final analysis success will depend more on the manner in which the commission performs its functions than on the specific provisions of the acts.

The over-the-counter markets. The over-the-counter markets constitute a very important part of the country's financial mechanism. The volume of business in over-the-counter transactions is impossible of accurate measurement, but it is thought to equal that of the organized exchanges, including transactions in both bonds and stocks. The large majority of all bond sales, even for those listed on the exchanges, takes place over the counter. The number

of over-the-counter houses has been estimated at 5,700, while the stock exchange members number fewer than 2,800.[1]

The regulatory problems presented by the over-the-counter markets are serious. On the one hand, to leave these transactions unregulated, while regulating the organized exchanges, would encourage their expansion at the expense of the better policed and more adequately regulated exchanges. Furthermore as regards the protection of investors it is more necessary to regulate such undisciplined markets than the more publicized exchanges. On the other hand, the administrative task of supervising activities of thousands of scattered dealers would require a cumbersome and expensive bureaucracy. Under the limited jurisdiction of the Act of 1934, the Securities Exchange Commission proceeded with the registration of brokers and dealers, having received 8,276 applications as of June 30, 1937, and having 6,736 registrations effective on that date; and it has also drafted rules to define the "manipulative, deceptive and other fraudulent practices" prohibited by the act.

Persuaded that further regulation was essential, the commission proposed and Congress enacted the Maloney bill, approved June 25, 1938, as an amendment to Section 15 of the Securities Exchange Act of 1934. This act provides a comprehensive program for coöperative regulation to be carried out largely through voluntary associations of investment bankers, brokers, and dealers. Over-the-counter dealers and brokers may organize into one or more national securities associations and affiliated associations. Such associations may apply for registration, filing registration statements analogous to those required of exchanges. The registration becomes effective if the commission finds that membership in the association is open to all who qualify, with fair representation for members, and that the rules are designed, not to permit unfair discrimination or fix minimum profits or commissions, but rather to protect investors and the public interest, to perfect the mechanism of a free and open market, and to provide for fair and orderly procedure in disciplining members. Furthermore the commission may review disciplinary action of such associations, suspend or revoke the registration of an association, and suspend or expel members of an association.

[1] *The Security Markets*, Twentieth Century Fund, 1935, pp. 221–222.

Changes in the rules and regulations of an association must be filed with the commission, which may set aside or amend any rule of an association. Though membership in such associations is voluntary, the provision that the rules of an association may require that no member shall deal with any non-member dealer except on the same terms and conditions as are accorded to the general public (except for granting one another dealers' discounts) is expected to make membership attractive. The enactment of the law was followed by considerable activity in the formation of over-the-counter associations.

Other corporation problems: The holding company. The holding company is under attack, not only because it provides an easy method of forming large combinations and so renders easier the securing of monopolistic control in a particular field, but more especially in recent years on the ground that it enables a corporation to conceal from the government or from the general public connections which might not find general approval and permits the corporation to conduct operations indirectly in which it could not lawfully engage directly. For example a national bank is not permitted to have a branch bank in another state, but there is nothing to prevent a holding company from buying up the controlling stock in two national banks in two adjoining states and thus engaging in what is called "chain banking." Prior to the banking act of 1933 there was no provision of law which authorized any examination of the condition of the holding company, because of the fiction that such a holding company was not engaged in the banking business. Many banks in the speculative era before 1930 were affiliated through holding companies with companies primarily engaged in dealing in securities, while some were affiliated with real estate and other investment companies. Granted that the relationships thus created were quite within the law, there is today little question of the unsoundness of such connections and of the undesirability of any device which conceals them. In the field of the public utilities — particularly of the light and power utilities — the use of the holding company has in certain instances permitted and encouraged the formation of vast, unwieldy combinations whose main reasons for existence have been either the opportunity afforded promoters to

utter large issues of speculative securities, to their own enrichment, or the possibility which intercompany financial transactions afforded of so raising the apparent cost of generating and distributing light and power as to justify public service commissions in sanctioning rates which were in fact excessive.

Federal control of holding companies. The growing conviction that the holding company requires regulation has found expression in a number of recent acts, of which the public utility holding company act, passed by Congress in August, 1935, is the most important. This act, which applies exclusively to electric and gas public utilities, defines a holding company as any company which directly or indirectly controls ten per cent or more of the voting stock of a public utility company or any person (individual or company) which directly or indirectly exercises such a controlling influence over the management or policies of a public utility or holding company as to make it appropriate that such persons be subject to the obligations, duties, and liabilities imposed on holding companies. Contrary to common belief, the act does not outlaw the holding company in these utilities, but it does direct the Securities and Exchange Commission to forbid, after January 1, 1938, the further existence of complicated public utility organizations which are predominantly interstate in character. It also confers on the commission rather broad powers of regulation with respect to the issue of securities, the acquisition of utility assets, intercompany financial transactions, reports of financial condition, and the like.

In general it is the purpose of the act to restrict each holding company system (holding company and its subsidiaries) to a "single integrated public-utility system and to such other businesses as are reasonably incidental or economically necessary or appropriate to the operation of such integrated public-utility system." By this is meant a system whose utility assets are either physically interconnected or capable of interconnection and which under normal conditions may be economically operated as a single coördinated system, confined in its operations to a single area or region not so large as to impair the advantages of localized management, efficient operation, and effective regulation. The commission has however certain discretionary powers to permit a holding company to control

more than one such system where peculiar circumstances seem to warrant it.

The act also directs the commission to require that necessary steps be taken to ensure an equitable distribution of voting power among the security holders of the holding company system and to see that the corporate structure of the system is not unduly or unnecessarily complicated by the continued existence of any company in the system. In carrying out this provision the commission shall require each holding company over which it has authority "to take such action as the commission shall find necessary in order that such holding company shall cease to be a holding company with respect to each of its subsidiary companies which itself has a subsidiary company which is a holding company."

With the detailed provisions of the act we cannot here concern ourselves, but it is important to note that the prime objectives of the act are to protect the interests of consumers and investors and to make possible more effective regulation by state commissions.

The trend towards control of the holding company is seen also in the banking act of 1933, which provides that both the holding companies which may control a member bank and the companies controlled by a member bank shall be brought directly under the supervision of the banking authorities. Similarly the emergency railroad transportation act of 1933 provides that where a holding company, not a carrier, is authorized to acquire control of two or more carriers it is to be considered as a carrier for the purposes of the act. As the reader will recognize, the task of the federal government in seeking to simplify corporate structures is complicated by the fact that corporation charters are granted by the individual states.

The SEC and the utility holding companies. The enforcement of the public utility holding company act of 1935 was at first seriously hampered by uncertainty as to its constitutionality. On March 28, 1938, the Supreme Court in its decision in *Electric Bond and Share Company* vs. *Securities and Exchange Commission* upheld the registration provisions of the statute and thereby cleared the way for the more active enforcement of the substantive provisions of the law. In June, 1938, the commission was able to announce that utility

holding companies having 99 per cent of the total assets of such companies had registered under the act. The registration of utility holding companies has supplied the commission and the public with important data not previously available. The commission's substantive regulation has been concerned with the development of uniform systems of accounts, supervising the issuance of securities, controlling the acquisition of physical assets or securities by holding companies, and regulation of affiliated service companies. Several holding companies have voluntarily taken steps for the simplification of their systems to comply with the requirement that holding companies be confined in their operations to a single integrated public utility, and all companies have been instructed to prepare such plans by the end of 1938 It has been the commission's general policy to enlist the constructive coöperation of the industry, and to this end it has endeavored to impose only the minimum of necessary restrictions.

Divorce of ownership and control. One of the most serious of all the problems connected with the corporate form of organization is that created by the growing separation of ownership and control. Control is an attribute of a certain type of stock, usually common stock, which carries with it the right to vote at meetings of stockholders for the election of directors and on matters of general policy The ownership of more than fifty per cent of the voting stock means practically complete legal control of the corporation and involves necessarily the submission of the minority stockholders to the rule of the majority. This situation, even though it means for some an effective separation of ownership and control, is ordinarily the natural one. Presumably the majority stockholder is active himself in the business and is as interested as anyone could be in its most efficient management, since the interests of majority and minority stockholders are generally not at variance with each other, the minority stockholders do not suffer.

Frequently however a minority interest is sufficient to give control where there is no single interest holding an actual majority. By securing sufficient proxies — a signed form by which a stockholder transfers the right to vote his stock at a particular meeting — a minority stockholder is sometimes able at the stockholders'

meeting to elect his own board of directors and to have his policies ratified. There are in addition various legal devices by which the owner of a small part of the total capital stock outstanding may sometimes secure actual control of the corporation. One such method, known as "pyramiding," is to arrange a series of holding companies with each company owning the majority of voting stock in the one lower down. Complete legal control can thus be secured by an investment of a small fraction of the value of the properties controlled. For example, the Van Sweringen brothers are said to have controlled eight Class I railroads having assets of over two billion dollars with an investment of less than twenty million dollars.[1] Another legal device, to which there was frequent resort in the past decade, is the issue of two types of common stock, non-voting and voting. In the reorganization of Dodge Brothers, Inc., effected in 1925, neither the preferred nor four-fifths of the common stockholders were entitled to vote in the election of directors. The possession of 250,001 shares of voting common stock, acquired at a cost of less than two and one-half million dollars, is said to have given Dillon, Read, and Company control over the one hundred and thirty million dollar corporation.[2] By other legal devices of this character a large majority of the stockholders of a corporation may be deprived of virtually all rights of control.

The type of control in which the divorce between ownership and control is most complete has been called "management control." In this case we find no single stockholder nor group of stockholders in possession of sufficient stock to give even minority control, and legal devices for securing control are either absent or ineffective. Nevertheless those who occupy the position of management — placed there perhaps at the time of the organization of the company — may use the power which management has over the proxy machinery so effectively as to make it easy to retain their position or to dictate their successors. Occasionally a powerful stockholder can rally around him sufficient opposition to throw out the existing management, as in the successful struggle of Mr. John D. Rockefeller, Jr., against the management of the Standard Oil Company of

[1] Berle and Means, *The Modern Corporation and Private Property*, 1933, p. 73.
[2] *Ibid.*, p. 75

Indiana in 1929, but it is an expensive business and rarely attempted.

As an example of management control we may cite the case of the Pennsylvania Railroad Company; in 1929 the largest stockholder owned 0.34 per cent of the total stock outstanding, the twenty largest stockholders together held but 2.7 per cent of the total stock, and no officer or director was to be found in the list of the twenty largest stockholders. In the American Telephone and Telegraph Company the same situation is found to exist, with no stockholder in 1928 holding as much as one per cent of the total stock outstanding, and with the twenty largest stockholders owning only 4.6 per cent. Likewise no single stockholder in the United States Steel Corporation held in 1928 as much as one per cent of the total stock outstanding, and the twenty largest holders owned but 6.4 per cent of the total amount.[1]

Berle and Means have found that in the two hundred largest non-banking corporations in the United States control seems to be divided as follows: [2]

	By number	*By wealth*
Management control	44%	58%
Legal device	21%	22%
Minority control	23%	14%
Majority control	5%	2%
Private ownership	6%	4%
In hands of receiver	1%	Negligible

The consequences of this divorce of ownership and control, of the weakening of the position of the common stockholder as contrasted with that of management, raise a problem of fundamental importance. As we have seen in Chapter VI, private capitalism proceeds on the assumption that the institution of private property furnishes in profits a compelling motive to the most effective use of the instruments of production and that in the long run society gains thereby. But in the case of many large corporations ". . . it is no longer the individual himself who uses his wealth. Those in control of that wealth and therefore in a position to secure industrial efficiency and produce profits are no longer, as owners, entitled to the

[1] Berle and Means, *The Modern Corporation and Private Property*, 1933, pp 84-87
[2] *Ibid*., p. 94.

bulk of such profits. Those who control the destinies of the typical modern corporation own so insignificant a fraction of the company's stock that the returns from running the corporation profitably accrue to them in only a very minor degree. The stockholders, on the other hand, to whom the profits of the corporation go, cannot be motivated by those profits to a more efficient use of the property, since they have surrendered all disposition of it to those in control of the enterprise. The explosion of the atom of property destroys the basis of the old assumption that the quest for profits will spur the owner of industrial property to its effective use. It consequently challenges the fundamental economic principle of individual initiative in industrial enterprise. It raises for re-examination the question of the motive force back of industry and the ends for which the modern corporation can or will be run." [1]

Concentration of control. A disquieting feature of present-day business is the extent to which the control of corporate organizations, especially of the large corporations, has tended to become concentrated in the hands of a relatively small number of business and financial leaders. The actual control of the business of a corporation is, as we have seen, vested in the management; that is, it rests in the hands of the directors and officers of the corporations. We have already noted that nearly half of all non-banking corporate wealth in the United States is vested in the 200 largest non-banking corporations. This large aggregate of wealth is thus immediately controlled by the 2,000 men (approximately) who compose the boards of directors of these corporations.

The director of a large corporation is rarely interested in that corporation alone; he may hold a dozen directorships in other corporations. For example in 1930 the directors of the United States Steel Corporation held 174 directorships in other companies, the directors of General Motors held 167, and those of the New York Central Railroad 306.[2] Interlocking directorates which result in monopoly control are forbidden by the Clayton act, and it is not suggested that these just mentioned do foster monopoly in the steel business or in the manufacture of motor cars or otherwise. The

[1] Berle and Means, *The Modern Corporation and Private Property*, 1933, p. 9.
[2] Corey, *The Decline of American Capitalism*, 1934, p. 391.

point is that they cultivate a community of interest among the most powerful business leaders of today which tends toward a common policy so far as concerns the relations of corporate business toward the government, labor, and the general public.

Financial control. Another aspect of this same problem is found in the increasing degree to which control of the large corporation is found to be lodged with financial leaders rather than with industrial leaders of the type, let us say, of Andrew Carnegie or Henry Ford. As a natural result of the financial ties existing between a corporation and its financial intermediary and the quite understandable desire of banks to protect investments which they have made and in part because of the desire of banking houses to assure themselves of the profits which come both from the ordinary financing of a corporation and from promotions and reorganizations, banks have acquired an increasing degree of representation on the boards of other corporations. Fifteen New York City banks are said to have held 1,762 interlocking directorships in other financial, industrial, and utility corporations in 1899 and 5,324 directorships in 1931, and even more striking evidence of this intimate relationship is the report that 167 individuals associated with J. P. Morgan and Company and allied banking houses held, in 1929, 2,450 interlocking directorships in corporations with assets of 74 billion dollars—22 per cent of the total assets of all corporations.[1] While these figures do not prove that banking control dominates industrial life, they do indicate that it has reached large proportions.

The corporation as a factor in modern economic life. The extension of the corporate form of organization into all fields of economic life, as well as the growing importance of the large corporate combinations, make the corporation an inevitable target for hostile criticism. In the preceding pages we have reviewed briefly some of the problems to which the corporate form of organization has given rise; certain important defects and disturbing tendencies have been disclosed, but it by no means follows that the corporation has degenerated into a menace to society.

It is scarcely too strong a statement to assert that there has been no organizational device in the past hundred years which has

[1] Corey, *op cit*, p 402

promoted so effectively the expansion of business activity as has the corporate form of organization. The corporation lies at the very basis of large-scale production, the intricate division of labor, the minute specialization in production, and the increased flow of low-priced goods of modern times. We think of these as the results of the Industrial Revolution and of the introduction of factory methods of production, but while large-scale production might have been possible, it is doubtful if it would have made much headway, without the corporate form of organization.

The utility of the corporate form has manifested itself chiefly through its influence in directing the savings of the multitude into the hands of the business men. The limitation of liability in case of insolvency has made investment in corporate stocks very attractive and has been an important means of gathering in savings from diverse sources. Formerly an investor had to assume risks all out of proportion to his knowledge of the particular business, and hoarding was often the only practical way open to the small investor to conserve his wealth for future use. Today he can invest in corporate securities either directly or indirectly through the medium of savings banks and insurance companies.

The small denominations in which shares of stock and bonds are issued and the ease with which they may be transferred facilitate investment in corporate securities. One or two or three men may shun an unproved but attractive venture, whereas a thousand men acting together may be drawn to it, for each one knows that the loss of his stake in the enterprise will not cripple him financially. He knows further that as a rule he may sell his interest at any time he chooses to one whose faith is greater and thus withdraw before suffering the total loss of his investment. Advance in industry and commerce is conditioned upon the taking of risk, and any device which so divides and scatters the risk as to make it less burdensome to the individual will in the long run be an aid to production.

In conclusion we should perhaps call attention to the fact that most of the defects of the corporation seem amenable to control. A corporation enjoys such powers only as are granted it by the state in which it is incorporated. If for example it were believed that the holding company was undesirable the remedy is at hand, if

it were felt that all common stock should carry with it the right to vote the state could so ordain, or if it should seem wise to try to give the minority stockholders more effective representation the state could provide for cumulative voting and say that a holder of a share of stock might if he chose cast ten votes for one director rather than one vote for each of ten directors. Unwholesome practices connected with the issue and sale of securities are certainly within the control of governmental agencies and can be eliminated. Abuses which spring, not from the specific form of organization, but from a low standard of business ethics offer a more formidable problem, but if one surveys the advances made in this regard from the period of the sixties, let us say, he will find sufficient ground for anticipating that there may be developed codes of business ethics in which the welfare of society may occupy a more prominent place than heretofore. The first step in this direction lies in a general agreement as to what constitutes the public interest and what practices affect it adversely. This we do not now have in sufficient measure even with respect to the problems discussed in this chapter.

XLIII

ECONOMICS OF RAILROAD TRANSPORTATION

The railroads and the public. In some countries railroad transportation is a government function. In the United States and many other countries however the railroads are owned and operated by corporations, whose stocks and bonds are the property of thousands of individual and corporate investors; they are private wealth, owned and operated for private gain.

The fact of private ownership and operation however does not prevent the railroads standing in a class apart from the ordinary types of business enterprise, with public relations of the utmost importance. On the one side the railroads are in a peculiar position of dependence upon the government. So great is the public interest in the development of transportation, that the government has in many ways come to the aid of the railways in their initial stages. Generous land grants were made by the United States government to several of the railroads, and in many cases state and local aid was forthcoming to assist in their building. What is of far greater importance, the government has given the railroads the advantage of its right of *eminent domain*, by virtue of which private property can be condemned for public purposes at a price determined by the courts or other authority. The railroads have to acquire their "right of way" by the purchase of land from private owners, and despite the fact that in the pioneer days of railroad building the advantages which might accrue from proximity to a railroad line were so highly esteemed that in many regions the right of way was donated and in some cases even the grading of the right of way was done gratuitously by the people of the community, the power to secure the condemnation of land was in many instances of inestimable advantage to the railroads in enabling them to secure the most desirable routes at reasonable prices. The public has in fact entered into a quasi-partnership with the railroads in these respects.

The dependence is not all on one side. The reason back of government aid is the supreme dependence of the people upon the railroads. A railroad is not a mere private business. The local department store can conduct its business about as it sees fit. It can sell what it chooses, refuse to sell certain things, or close up shop entirely, and it can charge whatever prices it can get. If the people do not like its goods or its prices, they are free to trade elsewhere. With the services of the railroads the situation is entirely different. The people are absolutely dependent upon the railroads — in most cases upon one railroad. This dependence in the United States has reached the point where a stoppage of service might assume the proportions of a national calamity and where inadequate service or extortionate rates might neutralize the advantages of large-scale production or at any rate impose a great handicap upon producers and consumers.

Public utilities. In return for its cooperation the government therefore imposes upon the railroads the obligation to give adequate service to the public. The right of way and equipment must be kept up according to certain standards, a certain minimum number of trains, passenger and freight, must be run, and schedules must be maintained at all times, in winter and summer, in good weather and bad. Safety devices must be installed, and finally the rates charged for railroad services must not be unreasonable. The public has become a partner in the railroad business, it has a vital stake in railroad services, and in return for the privileges granted it demands of the railroads adequate service and reasonable charges. This double relation of dependence and obligation between the railroads and the public is what puts them, along with certain other public service enterprises, such as the street railways, the telephone, the telegraph, water supply, etc., in a special class apart from private business in general, to which has been given the appropriate title of *public utilities* or *public service corporations.*

The railway problem. Experience has shown that the operation of economic forces, such for example as determine the quantity, the quality, and the prices of the products of competitive industry in general, has failed to bring corresponding results in the field of railway transportation. When competition has prevailed, it has

generally been of the cutthroat variety and has brought in its wake disturbing discriminations in rates, disorganization of service, and other abuses. On the other hand, competition has frequently broken down altogether, introducing serious problems of monopoly. The investment in the railroads is so great that their methods of financing have attained public importance. Differences between the railroads and their employees are matters of vital public interest, since railway strikes or lockouts may easily play havoc with the nation's commerce. There is thus a "railroad problem," whose public importance can scarcely be exaggerated. Though, as we have seen, it has many sides, the heart of the problem is in the determination of the rates charged by the railroads.

In our study of railway rates we shall find it convenient to posit the first steps of the inquiry on the assumption of free competition and the entire absence of governmental regulation, although the reader will doubtless recognize that this is not the situation existing today in the United States, whose railroad conditions are the result of more than sixty years of regulation both by state and federal authorities.

Railway expenses. The railroad business furnishes a most striking example of an industry subject to the law of decreasing costs. The manner in which this condition emerges in an industry where there is economy in large capital investment and in which the plant of maximum efficiency is a huge organization has been made sufficiently clear in an earlier chapter. The railroad business represents an extreme example of this.

Fixed costs. Interest. To begin with, the railroads require a large investment of capital. Great expense must be incurred to bring a railroad into being. The right of way must be secured and improved, tracks must be laid, stations for freight and passengers must be erected, expensive equipment in the way of rolling stock, signalling apparatus, etc., must be purchased or leased. Funds for these purposes are secured through the sale of railway bonds or other securities; and the interest on bonds and notes, which must be paid at definite rates, as well as the funds which the railroad may have agreed to set aside each year to provide for the eventual redemption of such securities, taxes and the rentals of leased lines,

form the greater part of the railroad's fixed charges. In harmony with the nature of fixed charges,[1] these expenses bear practically no relation to the volume of traffic. Given the facilities which are prerequisites to the satisfactory handling of any traffic, an increasing amount of goods and passengers may be hauled over the road without increasing the fixed charges, up to that point where it is necessary to enlarge the existing facilities — double-track the system, provide four tracks in place of two, build extra switching connections, enlarge the terminals, etc. — in order to take care of additional traffic. When this point is reached additions to the fixed charges may be expected, but again, once they have been incurred, they will not vary until it is again necessary to enlarge the plant.

It should be noted that the investment in a railroad is relatively permanent, in the sense that there is almost no alternative use to which the equipment can be put if the railroad proves to have been an unprofitable venture. A factory building can sometimes be transformed into a storage warehouse, or diverted to the making of another article, but when once rails have been laid they can be used only to provide transportation, with the sole alternative of sale as scrap iron, and there is almost no way of recovering the very great expense of building the right of way, with its extensive cuts, grades, bridges, and tunnels. As always, the fixed charges are recurring costs, which represent the hire of capital funds which have been invested in the past, and which continue to accrue whether the road be operated or not. If the road should become bankrupt and be declared unable to meet these obligations, they are simply shifted once and for all from the railroad to the bondholders.

Other fixed costs and variable costs. *Both of these costs, which* could at any time be avoided by the decision to go out of business, include the expenses for the maintenance of way and structures, the maintenance of equipment, the cost of conducting transportation, and general expenses which cannot readily be allocated to any of the other groups.

Most of the general expenses, of which the chief item is the cost of maintaining general offices for the railroad system, can be assigned to the group of "other fixed costs." This cost bears but little

[1] See Chapter X for an analysis of fixed costs and variable costs.

relation to the size of the operations. The salaries of the officials go on regardless of the fluctuations in the volume of traffic, and the officials can handle a large volume of business with little extra expense, save for the clerical force. The cost of additional clerks, or the saving through the dismissal of clerks, is a comparatively small item in the total expense.

The maintenance of the right of way includes rail and tie replacement and supervision. While the rail maintenance is closely connected with the speed, weight, and number of trains, there is little relationship between the cost of renewing ties and the amount of the traffic. Tie replacement is necessary because of exposure to the weather and would have to go on so long as any traffic is to be carried. Furthermore the expense of tie replacement forms a much greater fraction of the total expense of maintaining the right of way than does rail replacement. The cost of superintendence bears some relation to the volume of traffic, but it does not increase in proportion to increases of traffic.

The maintenance of equipment varies according to mileage, but not in proportion to the volume of goods carried. Full utilization of the carrying space of freight and passenger trains and of the tractive power of locomotives does not cause a proportional increase in the charges for repairs as compared with incomplete utilization. It is estimated that about one half of these expenses may be considered to be largely independent of the volume of traffic.

The cost of conducting transportation consists chiefly of the wages of the trainmen, fuel for the locomotives, and terminal charges. The wages and fuel charges vary directly with the train miles, but they do not vary in proportion, for all trains are not of the same length, nor are they equally heavily loaded. The wage cost for a full train is not much greater than for an empty train, and while the fuel cost increases with the load to be hauled, it does not vary proportionately. The terminal expenses vary with the volume of business, but bear no relationship to the train miles. Here again a large percentage of the expenses, perhaps one half, may be considered as independent of the volume of traffic.

This brief description indicates that not only are the interest charges of the railroad business especially heavy but that the other

fixed costs are also unusually large, leaving a comparatively small part of total railroad costs in the class of variable costs. It has been estimated that about one half of the combined other fixed costs and variable costs and two thirds of the total costs can be considered as constant.[1]

Decreasing costs. The presence of this large proportion of constant expense indicates that a railway operates under conditions of decreasing costs. Up to the point of maximum utilization, an increase in the number of passengers to be carried or in the amount of freight to be hauled does not cause a proportional increase in the total expenses. Hence the unit expense will decrease. If we assume that the rates remain constant, while the traffic increases, profits must show an increase which is proportionally greater than the increase in the volume of business, and the reverse would be true if the rates remained constant while the volume of traffic declined. This explains the fact, which has frequently been noticed, that a moderate growth in the volume of traffic, with no increase in the rates, is often enough to raise a railroad from financial embarrassment to the position of a highly profitable business enterprise.

This statement that the railroad is an industry of increasing returns may be given concreteness by an illustration. The figures in this illustration are obviously chosen arbitrarily, but they will serve to illuminate the point. Assume that a railroad has a capital stock of $100,000,000 and an income of $100,000,000, and that its interest amounts to $28,000,000, its other fixed charges to $33,000,000, and its variable expenses to $33,000,000.

Gross revenue		$100,000,000
Expenses		
Fixed charges, interest	$28,000,000	
Other fixed charges	33,000,000	
Variable expenses	33,000,000	
Total expenses		94,000,000
Profits (available for dividends)		$ 6,000,000

The railroad after meeting all its expenses has left the sum of $6,000,000 which equals 6 per cent on its capital stock.

Assume now a 10 per cent increase in traffic. The revenue, on the

[1] Eliot Jones, *Principles of Railway Transportation*, 1924, p. 78.

assumption that the rate is the same on the additional traffic as on the old, will be $110,000,000. The interest and other fixed charges will be the same as before, it being assumed that the existing plant is adequate to handle the additional business, but the variable expenses may be presumed to increase in the same porportion as the traffic. The account therefore stands as follows:

Gross revenue		$110,000,000
Expenses		
Fixed charges: interest	$28,000,000	
Other fixed charges	33,000,000	
Variable expenses $33,000,000 + 10% of $33,000,000	36,300,000	
Total expenses		97,300,000
Profits (available for dividends)		$ 12,700,000

The railroad, after meeting all its expenses, has left the sum of $12,700,000, which is equal to 12.7 per cent on its capital stock. An increase of 10 per cent in the traffic has thus increased the profits by over 100 per cent.[1]

Joint costs. Suppose it is desired to know the exact cost of shipping a ton of coal from Scranton to New York. We will assume that two freight trains, one consisting of twenty cars, the other of thirty cars of coal, are sent from Scranton to New York. It is quite possible to determine with a fair degree of accuracy the variable expenses of each train. There will be charges for fuel, wages for the trainmen, some wear and tear on the tracks and the rolling stock, and a few other expenses which might have been avoided entirely had the particular train not been run. These expenses will be greater for the second train than for the first, but they will not be fifty per cent greater, and if we were to determine the proportion of the variable expenses which should be borne by each ton of coal we would find the amount somewhat greater in the first case than in the second.

A third train of thirty cars is sent from Scranton to New York, but in this case it carries mixed freight — coal, cement, hardware, textiles, and groceries — and furthermore some of the cars are

[1] This illustration is adapted from Eliot Jones, *Principles of Railroad Transportation*, 1924, pp 78 79.

incompletely loaded. Again we can determine the variable costs for the train as a whole, but to distribute them among the different goods carried according to the expense caused by each individual shipment is another matter. The easy solution would be to calculate a flat rate according to weight or bulk, but it is quite clear that this does not accord with our desire to find out the expense of shipping each individual package. It costs more, for example, to handle a ton of small packages than a ton of coal. A car fully loaded with textiles costs little more to haul than one only half loaded, and a meticulous effort to allocate the variable costs according to the cost of rendering the service might lead to the surprising result that it cost more to carry some of the textiles than others on the same train.

An accurate apportionment of the variable costs presents enormous difficulties, but when we take the other expenses into consideration the case becomes hopeless. The fixed costs are incurred to provide equipment and personal services which are available, not only for the passenger and freight services, but for specialized services within these two great groups, which cannot be apportioned on any scientific basis. Such costs we call *joint costs*, meaning *costs which are incurred in the simultaneous rendering of two or more separate services and which can be apportioned only arbitrarily between the separate services or products.* With regard to the greater part of a railroad's expenses it is impossible to tell what portion should be assigned to passenger service and what portion to freight service, to say nothing of further assignment of costs within each of these services. It is just as impossible as to say, for example, what part of the cost of raising cotton should be attributed to the cotton fibre and what part to the cotton seed.

Railway rates: Peculiarity. In competitive business generally, cost of production has a predominant voice in the determination of price, and even where monopoly prevails cost generally at least sets a minimum below which price will not fall. In railroading, owing to the presence of joint costs, it is impossible to tell absolutely what any particular railroad service costs, although the additional out-of-pocket expense for affording an extra increment of service can be closely estimated. It is evident therefore that specific cost

cannot furnish the principle for determining rates. Any theory that rates "ought" to be adjusted to the cost of rendering specific services is untenable. In practice the railroad managers know what their total costs are, and they seek to meet these total costs and earn what profit they may from the traffic as a whole.

Charging what the traffic will bear. In the absence of any specific cost method of determining railway rates, the officials charged with the duty of rate making have recourse to a practical method, which has been given the somewhat repellent name of "charging what the traffic will bear." Some goods can stand a higher rate per ton than others. A ton of pig iron may be worth from $25 to $30, while a ton of silk cloth would be worth thousands of dollars. A freight rate per ton or per pound which would be prohibitively high if applied to pig iron might have little effect on the selling price of silks and consequently on the volume of traffic in silk. A high rate therefore is charged for carrying goods which combine high value and small bulk, while rates on cheap bulky goods are relatively low.

There are of course limits to the rates determined on the principle of what the traffic will bear. The upper limit is that rate at which shippers would not send the freight or passengers would not travel. The long run aspect of this should be emphasized, for a rate which would at the moment give the maximum return might be disastrous in the long run. It is not to the interest of the railroads either to cause local industries to wither up and die or to force their relocation in places where more reasonable rates are charged. It is this limit which varies so widely according to the value and the bulk of the goods. The lower limit is the rate at which the traffic would yield a revenue to the railroad no greater than the additional cost which it would occasion. At this minimum rate the traffic may not pay its share of the fixed costs. However so long as it pays enough to cover its own variable costs (estimated as closely as the traffic officials are able) and ever so little above, it is good business to take it. So the cheap bulky goods, such as coal, trap rock, etc., are carried for what is actually less than the average ton mile cost, while other parts of the service are counted upon to make up the interest and other fixed costs for the whole. In this way it is

planned that the total revenue shall cover the total expenses and yield as large a net revenue as possible.

Discrimination between goods. Discrimination between goods is thus frankly an essential part of the prevailing method of determining railway rates. The word "discrimination" has an ugly sound, and a connotation of injustice is apt to attach to it. But a moment's reflection will show that such discrimination is not only inevitable but is in the interest of the people as a whole.

It will be granted that no railroad would be built unless the promoters of the venture could, with some degree of confidence, count on securing a gross revenue sufficient to cover all costs and yield a profit. If they are forbidden to charge different rates for transporting different classes of goods, it may not be worth their while to build the road, for a rate high enough to secure the maximum contribution toward the constant expenses from the high-grade freight would preclude the possibility of shipping bulky freight, while a rate low enough to attract the low-grade freight would materially reduce the profitableness of carrying the other. In other words, discrimination between goods may be the only condition under which a railroad can be built, and however much aggrieved a shipper of textiles may feel because he pays a higher rate per hundred pounds than a shipper of wheat, he is still better off than if there were no railroad, as his patronage of the road demonstrates.

Suppose now that a railroad, already built and in operation, sees an opportunity to add a new class of traffic, but only at a rate considerably below the scale paid on its present business. The new business will add little or nothing to the fixed expenses of the road, while the revenue obtained will contribute something, though perhaps not much, over and above the variable cost added by the new traffic. This will be profitable to the railroad, and it does no injury to the former patrons of the road. In fact it is possible that the contribution of the low-grade freight toward the fixed costs may make it possible to charge lower rates on the high-grade freight than would otherwise be imperative. This may be cold comfort to the shipper if his rate is in fact not reduced, but it should at least show him that he is not being injured by the low-grade business and that

his complaint should be directed against his own high rate and not against the other low rate.

That taking on additional low-grade traffic at rates less than the prevailing scale is to the interest of the railroad with unused capacity is quite clear; a railroad operates at decreasing cost, and an increase in the volume of traffic, even of low-grade traffic, is greatly desired, for it means greater profits. As a matter of fact discriminations between goods started in this way. At first mostly passengers and high-grade freight were carried, but it was soon perceived that it was worth while to take the low-grade freight at a lower rate in order to spread the fixed costs over a larger volume of traffic. A railroad which has reached the point of maximum utilization does not have the same inducement to increase the volume of traffic. In fact it may find greater profit in raising the rates on the least profitable goods, thereby reducing the volume shipped, in order to have more carrying space for the more profitable freight, at least until such time as it appears wise to enlarge its facilities.

Railroad competition. It has been suggested earlier in this chapter that the railroads present the extreme case of an industry subject to the law of decreasing costs. The reader is already aware of the fact that the state of decreasing costs is an unstable one, in which there are no economic forces capable of fixing a definite price. Cutthroat competition is the result, continuing until equilibrium is attained through sufficient reduction in the number of competitors or through combination or monopoly. The railroads furnish the most instructive example of all this Let us consider a hypothetical illustration.

The X & Y Railroad runs between A and B and has a monopoly of the traffic. It carries, we will assume, 100,000,000 tons of freight at fifteen cents a ton, and has therefore a gross operating revenue of $15,000,000. Four million dollars cover the interest charges, $4,000,000 the other fixed expenses, and $4,000,000 (four cents per ton), the variable expenses, leaving $3,000,000 to be distributed to the stockholders in dividends. A new road, the M & N, is built between A and B to compete with the X & Y. It charges the same rate as the X & Y; *i.e.*, fifteen cents a ton, and naturally diverts some of the traffic from the X & Y — 25,000,000 tons, let us say.

Assuming that variable expenses decrease in proportion to the loss of freight traffic, this leaves the X & Y in the following position:

Gross revenue		$11,250,000
Expenses		
Fixed charges · interest	$4,000,000	
Other fixed costs	4,000,000	
Variable costs	3,000,000	
Total expenses		11,000,000
Profits (available for dividends)		$ 250,000

The M & N railroad in the meantime is not satisfied with the share of traffic which it is getting and reduces its rate to twelve cents a ton, thereby attracting so much of the traffic of the X & Y railroad that the latter is forced to meet the reduction. Even if we assume that the X & Y regains its former share of the traffic, *i.e.*, 75,000,000 tons, its gross revenue will be only $9,000,000, while its expenses are still $11,000,000. Not only have profits been wiped out entirely, but only $2,000,000 is available for paying the interest charges. The railroad is accepting traffic at a rate not sufficient to cover the total expenses of furnishing transportation, but it must continue to accept the traffic, since at least it yields something to apply to the fixed charges, and it is better to lose $2,000,000 a year than to lose the whole $4,000,000 of interest charges, which would be the result if operations were suspended. And the road will be especially eager in seeking more traffic even at this unremunerative rate. If it could, for example, secure 10,000,000 tons additional freight, it would receive an income of $1,200,000 and incur an additional expense of only $400,000, leaving $800,000 to be applied to the interest charges.

The rate war may continue until the X & Y Railroad has nothing at all with which to pay the bondholders, but this does not necessarily end the competition. The corporation is reorganized, the stockholders are wiped out, and the bondholders become stockholders. Inasmuch as the major portion of the interest charges has been eliminated, the road is in a better competitive condition than before, and the warfare continues until one side or the other decides that peace on any terms is to be preferred to continued struggle and

allows itself to be absorbed by its stronger rival. Cessation of hostilities may come before this point is reached, on the basis of an agreement to maintain rates and pool either the traffic or the profits. In either case the result is the same — the elimination of competition.

That this competition may be disastrous to the railroads is self-evident, but it does not follow that the public is the gainer thereby. The shippers of goods at the competitive points doubtless have their markets widened for them by the lower rates, but this may prove to be harmful even to them in the long run, for a long drawn out period of cutthroat competition may lead to an artificial location of industries and their establishment on a scale which cannot be maintained under a rate system which is reasonably remunerative. The rise in the rates at the end of the rate war is certain to affect some industries adversely and may more than counterbalance the gain during the period of low rates. Society as a whole is bound to lose on account of the dislocation of industry caused by unremunerative railroad rates.

Local discrimination: Description. There will always be discrimination in the rates charged different classes of goods under the principle of charging what the traffic will bear, as we have seen. Two other kinds of discrimination are virtually sure to appear where there is competition between railroads.

The first of these is local discrimination, which in the broadest sense of the term occurs whenever the rates between different points for a given class of goods are not proportional to the respective costs incurred. If the New York, New Haven & Hartford Railroad should charge the same rate for shipping a case of firearms from New Haven to Hartford, a distance of thirty-seven miles, as it charges from New Haven to New London, fifty-one miles, this would be a local discrimination. If it should charge more to Hartford than to New London, its action would be even more clearly discriminatory. Finally if the rate to Springfield were equal to or less than the rate to Hartford, which must be passed through on the way to Springfield, we should have the extreme case, in which the shorter haul is included in the longer.

All variations from a flat rate per ton-mile, or other unit, are not

evidences of local discrimination. The goods must be of the same kind; different rates for coal and shoes do not necessarily imply local discrimination. They must be shipped in the same volume; it is quite natural to make allowance for carload shipments as compared with less than carload shipments. Finally the conditions of transportation must be essentially the same; we should expect the rate charged on trains running through mountainous country to be higher than rates charged on similar shipments through level stretches.

Examples. Local discrimination has been very marked in the past and still exists to no inconsiderable extent, and the situation sketched above is by no means without parallel in railroad history. "On first-class goods, at one time, the rate from Louisville to Montgomery was $1.26 per hundredweight. On to Mobile, 180 miles further south, it was only ninety cents. In the same territory the rate on kerosene oil from Cincinnati at times has been three times as much to interior points as to New Orleans, three times as far. . . . The rate on rope from San Francisco to Independence, Kansas, is seventy-five cents; while the same goods are hauled on through Independence, Kansas, much farther, to Missouri River points for sixty cents per hundred pounds. Wichita, Kansas, complains that cotton piece goods from New York by way of Galveston are rated at $1.36, while by the same route to Kansas City, 225 miles longer haul by that route, the charge is only ninety-three cents."[1]

The reason for local discrimination. The chief reason and the sole justification for local discrimination are found in the situation created by competition with another railroad or with a carrier by water. Two roads run between A and B, as shown on the diagram on the right.

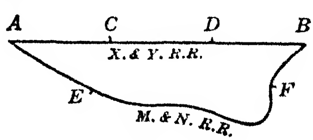

FIG. 6

The X & Y Railroad has a more direct route between A and B than does the M & N Railroad. The former charges, let us say, three cents a ton-mile on freight from A to C, C to D, D to B, and also on the stretch from A to B. The M & N Railroad likewise

[1] W. Z. Ripley, *Railroads, Rate and Regulation*, 1912, p. 216.

charges three cents a ton-mile on its intermediate traffic, but if it were to charge the same rate from A to B, the total charge per ton would be considerably greater than that from A to B via the X & Y. It therefore charges the same amount as the X & Y Railroad even though this charge may be less than that from A to F. Shippers and consumers at F complain against this discrimination; they say, with truth, that more is charged for the short haul than for the long haul, even though the short haul is included within the long haul, and they further maintain that this practice is unjust and should be stopped.

The defense of the railroad rests on the assumption that some of the through traffic is essential, either to justify the existence of a railroad through the intermediate towns or to enable the intermediate rates to be as moderate as they are. The reasoning underlying this is essentially similar to that used to explain discrimination between goods. The rate on the through traffic is remunerative in the sense that it covers the variable expenses involved and makes some contribution toward the constant expenses. Without the through traffic this amount would have to come from the intermediate traffic alone, and the rates would therefore be higher than they now are, or in anticipation of such a situation the road would not have been built at all. The discrimination, though real, is nevertheless essential to providing service for the intermediate points, and while the producers, shippers, and consumers at the intermediate points may be handicapped as compared with those at A and B, they are better off than they would be without the service or with still higher rates.

All discrimination between places cannot plead this excuse. Local discrimination is sometimes used to conceal personal discrimination in the interest of favored shippers, as will be noted later. Discriminatory rates to build up the land values of regions in which the railroad or its officials own land have not been unknown, and for other reasons the railroad officials may arbitrarily favor one place over another.

Conclusion. When two railroads are in competition for the through traffic, there will almost certainly be local discrimination on at least one road, and there is very likely to be discrimination

against local points on both roads. In our example we assumed that no discrimination was found on the X & Y Railroad, but as a matter of fact, this situation is scarcely to be expected, for this road will be tempted to try to regain some of the traffic lost to the M & N Railroad, by lowering its through rates. In fact if the route by the X & Y Railroad is very considerably shorter than the other, the former may be able to lower the rate to a point which will be absolutely unremunerative for the M & N Railroad and retain all the through traffic.

Discrimination between places is a serious thing. Towns against which there is discrimination are injured, their industries may be destroyed, and their population driven elsewhere. The towns which enjoy the low competitive rates are given artificial advantages over other towns which may offset or completely eliminate real natural advantages possessed by the latter. Population is attracted to the favored towns and the problems of city life intensified. Nor do the railroads themselves, as a whole, gain by the practice, for while a certain increase in traffic is likely to be stimulated by the lower rates, it is questionable whether it will be sufficient to compensate for the rate reduction. Nevertheless local discrimination is an inevitable feature of railroad competition, and it is probable that there would be a certain amount even under monopoly.

Personal discrimination. We have now to consider a third form of discrimination in railroad rates; namely, discrimination between persons. Such discrimination is peculiarly the result of railroad competition. Each road is forced to meet the cuts in the rates initiated by its competitors in order to get any business; if it can secure additional business by a secret contract with a particularly prominent shipper it is to its advantage, temporarily at least, to do so. Referring to our hypothetical example above, suppose that the X & Y Railroad quotes a secret rate of ten cents a ton to a favored shipper at a time when the published rate is still twelve cents a ton and thereby diverts 10,000 tons of freight from the M & N. It will secure $1,000 of additional revenue at an expense of $400, and its rival will find its revenue reduced by $1,200 and its expenses reduced by only $400.

The essential feature of personal discrimination is to select some

one shipper and attach him so firmly to the railroad by favors of one sort or another that he will patronize that railroad to the exclusion of all others. The tendency is quite naturally to favor the growth of industrial monopoly by restricting the area and scope of competition. Quite aside from its general social bearing, the growth of monopoly power among its shippers is not without danger to the railroad itself. When the fierceness of competition has died down and the railroad wants to readjust its rates, it may find itself no longer the master of the situation. The monopoly, ungrateful for past bounties, is in position to demand special rates by threatening to throw all of its business to the competing railroads.

Personal discriminations of all kinds — some of them straight rebates, and others concealed more or less successfully in allowances for demurrage, etc. — have been notorious in American railroad history. One of the outstanding examples is seen in the arrangement between the Standard Oil Company and the Cincinnati & Marietta Railroad. On shipments from Macksburg, Ohio, to Marietta the Standard Oil Company at one time not only paid but ten cents per barrel of crude oil, whereas the independent companies were charged thirty-five cents, but, more than that, the railroad turned over to the Standard Oil Company twenty-five cents out of each thirty-five cents paid by its competitors! In all fairness to the Standard Oil Companies of the present time it should be emphasized that this occurred many years ago and that the Standard Oil Companies today are among the most scrupulous in their observance of the literal terms of the printed tariffs.

Local discrimination is sometimes used as a means of concealing personal discriminations. An influential manufacturer may represent the only important manufacturing interests in a small city; if open but discriminating rates are quoted on shipments from that city his competitive power is obviously strengthened. Of course it is possible for other manufacturers to settle in the city and secure the same rates, but the process is necessarily a slow one.

Other examples might be cited, but they are not necessary to demonstrate how totally at variance with the ideal of the public utility is discrimination between shippers. It is quite contrary to all our notions of fair play that one man should be singled out and

favored to such an extent that others are ruined and driven out of business, particularly by an agency upon which the whole public is dependent and which has received important privileges from the public. The encouragement to monopoly which results cannot be viewed with complacence, nor can it be urged, with any cogency, that the railroads act in accord with the public interest in simply hastening the process of natural selection among industrial leaders, for there is little evidence to show that efficiency is apt to be a consideration of any importance in selecting men or companies as the recipients of railway favor. That the public guards its own best interest in demanding that public authority put an end to these abuses is scarcely open to question.

The breakdown of competition. A picture of the many undoubted abuses and the many practices at least open to serious question which follow in the wake of unregulated railroad competition is now before us. It is perfectly clear that competition fails in the railroad business to do what it accomplishes on the whole in industry in general; namely, to ensure the production of the commodities and services which the people want at prices at least not far from cost of production. In the early days of railroading competition was relied upon to protect the public against the monopolistic practices of the railroads, and early American legislation attempted to compel the maintenance of competitive conditions. This child-like faith in competition was woefully misplaced.

Not only does competition fail to produce the results hoped of it, while producing other results not foreseen or desired; competition proves itself either a pure figment of the imagination or at most an unstable and temporary situation, with an inherent tendency toward monopoly. Complete competition between railroads has probably never existed to any important degree; even under what would be called competitive conditions, a partial monopoly usually is found. If one examines the routes taken by the roads running between New York and Chicago he will realize that, while competitive rates exist on the through traffic, each of the roads has a more or less complete monopoly of its local traffic. In fixing the rates on local traffic the zeal of a railroad to secure monopoly profits is tempered only by the realization of the fact that the producers of

its region have to meet in the common markets the competition of producers of other regions and that therefore they cannot charge rates so high as to make it impossible for the local shippers to meet this competition.

The most important reason for the failure of competition however is that competition is self-destructive, it degenerates into cutthroat competition and ends in combination or monopoly. As we have seen, the presence of so large a percentage of expenses incurred for all services makes it impossible to equate the service to the expense of rendering it, and the fact that so large a part of the expenses are constant in the face of either a decrease or an increase in the volume of traffic places a premium on securing a large volume of traffic, up to the point of complete utilization of the plant. The result is that, when keen competition starts, the prime object of traffic managers is to secure traffic almost regardless of the rate. Rates will be cut, and they may fall so low as to cover only the variable expenses; secret rebates are given to attract shippers, while at the same time the shippers at the non-competitive points are charged monopoly prices. The end of this is the elimination of competition through consolidation or combination.

When once a road has gained a monopoly position, it is easy to hold. Even though the road may be securing monopoly profits which would ordinarily be more than ample to attract the investment of capital, it may be impossible to interest capitalists in the building of a competing railroad. Investors want to be very certain that the traffic is sufficiently large to assure both roads of an adequate rate of profit at the competitive rate on freight and passengers. The investment in a railroad is large, and when once made it cannot be transferred to other uses. Competition therefore gets under way very slowly and cautiously even in the presence of monopoly profits.

Monopoly inevitable and economical. There can be no escaping the conclusion that, on the whole and in the long run, monopoly in the railroad business is inevitable. And an equally important truth is that railroad monopoly is more economical than competition. Not only does a lower unit cost result from the more complete utilization of existing facilities, but when additional facilities must

be provided it is better to enlarge the existing plant than to build a new railroad. The laying of additional rails, the expansion of terminals, etc. entail, it is true, a considerable increase in the fixed charges but the additional cost is much smaller than that necessary to develop the right of way, lay the rails, and provide the equipment necessary for a new railroad. In other words there are economies to be obtained from large-scale operation. A fully utilized two-track road operates at a lower unit cost than a similar one-track road, and the same holds true of a four-track road as compared with a two-track road. Monopoly allows us therefore to reap the benefits to be derived from complete utilization of the plant and to realize the economies which result from large-scale operation. It represents a more economical use of the limited supplies of labor and capital.

Regulated monopoly. While it is pretty clear that monopoly of rail facilities is the condition best adapted to the public interest from the standpoint of the most efficient utilization of labor and capital, unregulated monopoly is the cause, not merely of fears which may be somewhat irrational, but of misgivings which are well grounded on facts and principles. Unregulated monopoly carries with it no guarantee of adequate service at rates which are only reasonably remunerative or of impartial treatment of all shippers. In fact we may expect the railroads to consider profits before service, to be influenced by private motives rather than motives of public service, to forget that they have, by reason of their quasi-partnership with the public, obligations which should be fulfilled.

From a practical point of view unregulated monopoly of the railroads is inconceivable in the United States. The choice lies, therefore, between (1) government regulation of the activities of private companies and (2) government ownership and operation. The problems arising out of this conclusion form the subject matter of the following chapter.

EXERCISES

1. It has been estimated that about 60 per cent of a railroad's expenses are independent of the volume of traffic, while 40 per cent are dependent on the volume of traffic. Assume the following conditions of cost and return for a railroad concern:

Operating income	$102,000,000
Fixed costs	60,000,000
Variable costs	40,000,000
Net income	2,000,000

(*a*) How would a 10 per cent increase in traffic affect the railroad's net income?

(*b*) How would a 10 per cent decrease in traffic affect the railroad's net income?

2. Of what importance is the distinction between fixed and variable costs to the problem of railroad rates? If a railroad between Pittsburgh and Chicago is already in existence and trains are running, what is the lowest rate at which the railroad could afford to accept additional traffic? Would it be feasible for the railroad to accept all traffic at such rates?

3. How does the presence of joint cost affect the determination of railroad rates? Why do not railroads charge uniform ton-mile rates?

4. A railroad runs from New Orleans through El Paso to Los Angeles. Is there any justification for charging a lower rate per ton for shipments from New Orleans to Los Angeles than for shipments from El Paso to Los Angeles?

XLIV

GOVERNMENT REGULATION OF RAILROADS IN THE UNITED STATES

The whole problem of transportation presents so large a subject that its complete treatment is quite beyond the limits of a general book such as this. Having obtained from the preceding chapter a broad concept of the economic nature of railway transportation, we shall have to be content with a study in the present chapter of what has been accomplished in the United States alone in the field of government regulation of privately owned railroads.

State regulation of rates: Early history. The need for regulation of the railroads was not clearly recognized in the pioneer days of railroad building and indeed did not find expression on the statute books until the years around 1870. In the preceding decades the primary need was felt to be the building of roads; competition could be relied upon to regulate them. Gradually however abuses sprang up which the public refused longer to tolerate, and legislation was passed in many states to control, to a greater or less extent, the activities of the railroads. It is significant of the spirit of the times that the individual states made the first move in this direction. The state of Massachusetts created its railroad commission in 1869, whereas no attempt was made by the federal government to assert effective control over the railroads until 1887, when it was virtually forced to enact legislation because of the states' lack of jurisdiction over interstate traffic.

The Massachusetts commission was of the supervisory type; it had no powers of compulsion but made investigations and recommendations to the railroads, relying upon the force of educated public opinion to enforce its decisions. Even with these decidedly limited powers it attained a considerable degree of success, and it was of importance as a pioneer and an example for other states to follow. It was not however the prototype of the railway com-

missions of today, which are usually clothed with mandatory powers.

The movement for strict control of railroads found its origin in the agricultural states of the Middle West, which in the early seventies were suffering from an overexpansion of agriculture, as well as from the universal depression in business. The chief complaint against the railroads was that exorbitant rates were charged on shipments of agricultural produce, and when the competitive wars of the seventies somewhat reduced the force of this argument the charge of local discrimination took its place. There was moreover a profound conviction that the common attitude of the railroad executives, who appeared to regard themselves solely as the representatives of private interests, without obligation to the public and not amenable to the legislative power, was dangerous and needed curbing The corruption evidenced in giving passes to those whose official position might enable them to lend support, in bribing legislatures, and in the skilful juggling of the finances of the roads for the benefit of insiders added to the force of the movement.

The "Granger laws." Early in the seventies many of the states of the Middle West passed the "Granger laws" (so called because of the support which the National Grange of the Patrons of Husbandry lent to the movement) forbidding discriminations and prescribing maximum rates which might be charged. Naturally enough the railroads at once contested the constitutionality of laws which threatened to curb their arbitrary powers We cannot enter into all the details of these cases, but we may note that the Supreme Court upheld the doctrine that the railroads were common carriers affected with a public interest which it was the inherent right of the state to safeguard and that it upheld the contention of the states that they had the right to prescribe maximum rates. In the course of time this absolute right to prescribe rates was modified to the extent of giving the railroads the right to appeal to the courts for a review of rates which they considered unjust.

Many of the Granger laws were ephemeral. Conceived in a spirit of bitterness and even of hostility to the railroads, they overshot the mark and reacted adversely on the interests of the railroads and of the public. Their chief services lay in bringing about

an enunciation of first principles, in blazing a path in the unknown and untried field of railway legislation, and in opening the eyes of the railroad executives to the possibilities of the future.

The conflict of jurisdiction. Conflict between state and federal authorities over the control of the railroads was inevitable as soon as the railroads emerged from that condition where they consisted of short disconnected lines only and became increasingly important as the carriers of commerce between the several states. The United States constitution gives to the federal government control over foreign and interstate commerce, while each individual state has authority over the commerce originating and having its entire course and destination within its boundaries. The difficulty of fixing rates for intrastate trade without such action having an influence on the interstate rates, or the reverse, is quite apparent.

In the Granger cases it was decided that, in the absence of federal legislation on the subject, the state had the power to prescribe rates to regulate intrastate traffic, even though such action might indirectly affect interstate commerce. But the Wabash case, decided in 1886, established the principle that, even though the federal government did not choose to exercise its power, it nevertheless possessed sole authority to regulate interstate commerce. Since it was inconceivable that interstate commerce should be left to the arbitrary control of the railroads, particularly since about three quarters of the traffic of the roads in 1886 was between the several states, Congress was forced to enact its first comprehensive regulatory legislation — the act to regulate commerce of 1887.

With two authorities determining rates within a state — the federal authorities in control of interstate rates and the state authorities in control of intrastate rates — the question of the dominant authority had to be settled. In case of conflict should intrastate or interstate rates rule? In the Shreveport Rate Case (1914) the complaint was made that the rates prescribed by the Texas Railway Commission on intrastate traffic constituted undue discrimination against traffic coming into the state from Shreveport, Louisiana. The court, in its review of the case, sustained the findings of the Interstate Commerce Commission that the state rates were unduly prejudicial to interstate commerce originat-

ing at Shreveport. It also sustained the commission's fixation of reasonable maximum interstate rates. Based on these two sets of items, the court required the railroads to eliminate the undue prejudice found to exist. This they were at liberty to do by lowering the interstate rates to the level of the state rates. They were also expressly permitted to eliminate the undue prejudice by observing on both classes of traffic the interstate rates set by the commission, even though this involved the disregard of the Texas state rates. The Shreveport case marks therefore a step toward more complete control over rates by federal authority. The final step in this direction, short of complete control over intrastate rates, was taken in the transportation act of 1920, which states that where a state-made rate discriminates against interstate traffic, the Interstate Commerce Commission shall prescribe a rate to remove this discrimination, which shall be binding, "the law of any State or the decision or order of any State authority to the contrary notwithstanding."

Federal regulation: The interstate commerce act of 1887. This act was made applicable to interstate and foreign freight and passenger traffic by rail or by rail and water; intrastate commerce was specifically excluded. Rates were to be just and reasonable; and personal, local, or commodity discriminations of an unreasonable character were declared unlawful. Each road must provide reasonable and equal facilities for the interchange of traffic with other lines, and it might not discriminate in rates between such connecting lines.

Charging, under substantially similar circumstances, more for the transportation of passengers or like kind of property for a shorter than for a longer distance over the same line in the same direction, the shorter being included within the longer distance, was declared unlawful. It was however provided that the Interstate Commerce Commission, created by this same act, might after investigation afford the carriers relief from this provision if special conditions seemed to warrant it. An attempt was made to ensure the continuance of competitive conditions by making unlawful any contracts or agreements between railroads to pool freight or to divide earnings. Rates and fares were to be printed and posted in

every station and filed with the Interstate Commerce Commission. Rates not so posted were unlawful. No advance in rates might be made without ten days' notice, but a reduction might be made at any time, provided immediate notification were given the public. Damages might be recovered from a railroad by a person injured through unlawful practices, and provision was made for a fine of not more than $5,000 for each offense upon conviction in the courts. Two years later imprisonment was made an alternative or additional penalty for certain offenses.

Finally the Interstate Commerce Commission was created as a supervisory and regulatory body. Under the law as amended to the present, the commission consists of eleven members appointed by the President by and with the advice and consent of the Senate for terms of seven years. The commissioners are forbidden to engage in any other activity, and no commissioner may own stocks or bonds in any railroad company. The commission has the right of investigation and report, either on its own motion or on the complaint of a shipper. It can compel witnesses to testify and to produce books and papers, it can order unlawful practices to cease, and if necessary it can apply to the federal courts to issue an enforcing order to the offending road.

Defects of the interstate commerce act. The interstate commerce act did not work out as anticipated by its framers. Many defects in the law itself were discovered. It appeared, after a series of court decisions, that the law had not given the commission adequate power to compel reluctant witnesses to testify. Although there was additional legislation in 1893, it was not until 1896 that the right of the commission to compel testimony and the production of evidence was fully established by a decision of the Supreme Court.

Another grave defect was that the commission's orders were not binding upon the carriers. They might obey them or not as they chose, and the only recourse was to take the case to the courts. If the courts were satisfied that a lawful order had been disobeyed, they might issue an injunction enforcing the order. But the delay involved in instituting such suits and in carrying an appeal to a higher court, as well as the fact that the courts allowed the railroads

to submit evidence which had not been previously submitted to the commission in its own investigation, made this method of regulation highly unsatisfactory.

The final blow came in 1897, when it was decided that the interstate commerce act did not give the commission the right to determine maximum rates. Henceforth it could, subject to review of its action by the courts, determine whether a particular rate was reasonable or not, but if it condemned a rate as unreasonable it could not fix the maximum that might be charged. A railroad could, by a series of minute rate shadings, postpone almost indefinitely the reduction of a rate to that point considered by the commission to be just. Shippers could still appeal to the courts against unreasonable charges, but this was a right which in the nature of the case would seldom be invoked.

Development of regulation to the period of war control. Judicial definition of the legal limits to the powers of the commission and a widespread feeling that more rigorous control of the railroads was desirable led to the passage of supplementary legislation, and beginning in 1903 the act to regulate commerce of 1887 has been frequently amended, either to attain the objects of the act or to bring the railroads under greater control than originally contemplated. The power to determine and prescribe maximum rates in the place of existing rates which it considered unreasonable was explicitly given the commission. It might moreover prescribe maximum joint rates on through routes and settle the division of the rates between the several carriers concerned.

The railroads' power to alter rates was circumscribed, first by requiring thirty days' notification of any alteration in the published rates and later by giving the commission authority to suspend the operation of any new rate, classification, regulation, or practice for a period of 120 days, with an additional period if necessary, while it made an investigation of the need or justice of the proposed change.

Equally important were the improvements in the procedure devised to carry out the orders of the commission. In order to decrease the time required to settle cases an expediting act was passed, giving precedence in the Circuit Courts to suits in equity arising under the interstate commerce act or the anti-trust acts

in which the United States appeared as complainant or where the Attorney-General certified that the public interest was involved. Appeals from the decisions of the Circuit Court were to be taken directly to the Supreme Court within sixty days after the decree had been given by the lower court.

Orders of the commission, except for the payment of money, were made effective within thirty days after the date of issue and remained in force for two years unless suspended by the commission or the courts. The railroads were debarred from pursuing their former tactics of obstruction by the imposition of a fine for neglect to obey an order of the commission amounting to $5,000 for each offense or each day of the same offense. In case of failure to obey the orders of the commission the Circuit Courts were furthermore directed to enforce compliance.

Protection was afforded the carrier against arbitrary or unreasonable action of the commission, it had the right to request the court to suspend an order of the commission and to appeal the decision of the Circuit Court directly to the Supreme Court. In no case however was an order of the commission to be suspended except on hearing after not less than five days' notice had been given the commission. Since 1906 the courts have contented themselves with reviewing questions of law and have considered the commission as a competent and effective administrative board, whose findings of facts and whose orders should not be questioned so long as they are confined within its lawful field. This has ended the troublesome question of admitting fresh evidence in the court hearing, a practice which not only prolonged the settlement but made conflict of decisions between the commission and the courts a frequent occurrence.

The provisions intended to eliminate personal discriminations were strengthened, in the first place, by making the published tariff the test of lawfulness, so that a rate other than the published tariff was *ipso facto* illegal regardless of whether other rates had been charged or not. Previously, in order to prove the existence of discrimination, the commission had found it necessary to show that a secret and preferential rate had been given in some specific case and that higher charges had actually been levied on like and con-

temporaneous shipments. In the second place, the penalties for the offer, solicitation, or acceptance of rebates were made much more stringent than formerly. The commission was also given greater authority over long and short haul charges in order to prevent unjustifiable discriminations.

Maintenance of competition. The maintenance of competitive conditions on the railroads has been a cardinal feature of American policy. We have already noted the legislation against pooling. Pools were formed at quite an early date to avoid the ruinous effects of cutthroat competition, and some of them were quite successful in spite of the fact that they had no standing before the law. According to the common law doctrine, agreements or contracts in restraint of trade — and pools to divide earnings or freight were open to this interpretation - are null and void and cannot be enforced by court action. Declared illegal and made punishable under the federal law, pooling gave way to traffic associations designed to control and maintain rates and to combinations or consolidations of competing lines. To meet this situation the government invoked the authority of the Sherman anti-trust act of 1890, which declared illegal "every contract, combination in the form of trust or otherwise, or conspiracy in restraint of trade or commerce among the several states, or with foreign nations," and in a number of notable cases succeeded in forcing the dissolution of the combinations.

In recent years however the fear of a railroad monopoly has somewhat died down. There is perhaps less to be gained from competition between railroads, as the decrease in the unused capacity of the roads has in large part removed the incentive to cut rates in order to secure traffic. Certainly the more effective control exercised by the Interstate Commerce Commission has contributed to this change of attitude, as well as a somewhat belated appreciation of the inherent difficulties of maintaining competitive conditions and of the desirability of securing the advantages which may come from unified control and elimination of needless duplication in facilities and services.

Limitation of rate increases. One effect of the regulation of rates, unforseen when the legislation was passed, worked great

injury to the railroads. Most of the legislation came during the latter years of the nineteenth century, when the general price level was, as we have seen,[1] at its lowest. After that there came an extraordinary increase in prices in general. Supplies and equipment used by the railroads increased in price, and the higher costs of living made it necessary to grant wage increases. Railway income did not increase proportionately because of the legal restrictions upon rates. However distressing this situation was to the railroad managers, it was generally viewed with equanimity by the public and by the Interstate Commerce Commission, which refused to sanction rate increases sufficient to cover the extra costs. The result was disastrous. Railway profits declined and were in many cases succeeded by losses. Investors were reluctant to put their capital into railroad securities, their timidity being increased unfortunately as a result of serious financial mismanagement of some of the more prominent roads. The consequence was an impairment of credit which made it impossible to procure funds for needed improvements, extensions, and equipment.

War-time operation by the government. From December 28, 1917, to March 1, 1920, the national government assumed control of the operation of the steam railroads and of services ancillary to their operation, such as the express business, sleeping and parlor cars, telephone and telegraph lines, etc. The federal control act, under the authority of which operations were carried on, provided for compensation to the railroad owners for the period of government operation and for maintenance, repairs, and renewals, and it gave to the President broad powers to alter rates, classifications, regulations, and practices.

The reason for this step was the manifest inability of the railroads to handle the enormous volume of traffic with the dispatch necessary in war times. In part this was the result of the impaired credit and rail facilities which we have described above, but in the main it came from a lack of unity in railroad management necessary to meet a great crisis. Under federal operation the railroads were unified and their physical resources — terminals, equipment, and

[1] See Chapter XXIII for an account of the changes in the general level of prices since the Civil War.

repair shops — were pooled. Economies were effected in operation through consolidated ticket offices, centralized purchasing, and standardization of equipment. To handle labor, which was both discontented and inadequate to the needs, the Director General of Railroads set up a Board of Railroad Wages and Working Conditions, extended the eight-hour day throughout the railroad service, raised wages, and standardized them throughout the United States. The increased cost of operation was offset in part by increases in freight and passenger rates, both interstate and intrastate.

The essential purpose of the federal administration was to keep men and materials moving smoothly and rapidly to their destinations. This was accomplished successfully. From the standpoint of financial operations, on the other hand, the Railroad Administration was not successful. Huge deficits were piled up, approximating $1,200,000,000 for the twenty-six months of federal operation.

Return to private operation. In May, 1919, President Wilson announced that the railroads would be returned to their owners on January first, 1920. The necessity of finding a more acceptable solution to the railroad problem than had hitherto been attained and the widespread interest of the general public in the question led to a careful survey of the possible conditions under which the roads should be operated in the future. In spite of the financial tangle in which the government had involved the railroads and the natural reaction of the public to the many inconveniences of war-time control, the possibility of continuing government operation was not immediately discarded. In fact a certain portion of the people wanted the ownership of the railroads, as well as their operation, vested in the government. Railroad labor, which had found the government a more sympathetic master than the private companies, was especially insistent on this and to this end worked out a fairly detailed plan, commonly known as the Plumb Plan. But the advocates of this plan were not numerous and influential enough to win Congress to further experimentation in government operation.

The transportation act of 1920. This law, known also as the Esch-Cummins act, was passed as an amendment to the act to regulate commerce of 1887. It represents an honest and sincere

effort on the part of its framers to provide a solution of the railway problem which should be just to the owners of railways, to railway labor, and to the general public. While the act has been amended in several important particulars, many of its provisions are still in force, and in any case it warrants study as marking a turning point in the theory of railroad regulation.

Financial provisions. It was clearly recognized that rates must be adjusted so as to give the railroads revenue adequate to attract investment in their securities. At the same time it was realized that a general rate schedule which would assure this income to the majority of roads would yield to some a revenue far in excess of their needs. The problem was resolved as follows: the commission was empowered to divide the railroads of the country into groups and to establish rates which would allow each group as a whole to earn from five and one half to six per cent on the value of their property used in transportation. The law thus determined what was to be considered a fair return on the property; but after March 1, 1922, the commission was given broad powers of determining a fair return, taking into consideration the transportation needs of the country and the necessity of expanding railroad facilities. A rate of five and three quarters per cent on the value of the property was fixed as constituting a fair return. One of the most difficult problems, both of theory and of practice, which confronts a commission charged with determining reasonable rates for any public utility, whether it be a railroad or an electric light and power company, revolves around the determination of the "value of the property." Discussion of this problem however lies beyond the scope of this text.[1]

It is worth emphasizing that these provisions, which have been changed by recent legislation, did not in any sense guarantee a fixed rate of return to a given railroad. A strong railroad might earn from ten to twenty per cent on the value of its property, and a weak railroad might have no net earnings or even show a loss. It was expected however that the average earnings would be as

[1] The reader will find an excellent treatment of this topic in Slichter *Modern Economic Society*, 1931 Chapter XVIII, and in the special books on the subject listed at the end of this section.

high as five and three quarters per cent, though in the period since the enactment of the law this expectation has not been realized.

In order to guard against excessive earnings the so-called "recapture clause" was inserted. A railroad which in any year earned more than six per cent on the value of railroad property used in transportation was required to turn over one half of the excess to the government. The government's share went to a revolving fund to be used either (1) for loans to needy railroads to meet expenditures for capital account or (2) to refund maturing securities originally issued for capital account or to purchase equipment to lease to the railroads.

As might be expected the railroads did not accept without protest the claim of the government to share in their earnings, but in a case eventually carried to the Supreme Court, the Dayton-Goose Creek Railway *vs.* United States, the constitutionality of this provision was finally settled to the satisfaction of the government The unsatisfactory financial showing of the railroads made it however little more than a cause of irritation to the railroads, and in 1933 this provision was repealed retroactively.

Several other provisions of this act amplify the control of the commission over rates in other respects than implied above. The commission is specifically given power to prescribe minimum rates. The importance of this provision lies in the recognition of the desirability of calling a halt to cutthroat competition. The commission may therefore refuse to allow a railroad to set a rate which would be manifestly unfair to a competing road. Mention has already been made of its power to establish intrastate rates when the existing state rates unjustly discriminate against interstate or foreign commerce. And lastly the commission's power to relieve a carrier from the rigid interpretation of the long and short haul rule was somewhat circumscribed by the proviso that the commission shall not allow the establishment of any rate which is not reasonably compensatory for the service rendered.

Railway consolidation. In an earlier section of this chapter mention was made of the change in public policy with respect to the maintenance of competition between railroads. This is nowhere more clearly illustrated than in the section of the transportation act

of 1920 dealing with consolidations. The Interstate Commerce Commission was directed to draw up a plan for the consolidation of the railroads into a number of large systems. It was intended to preserve competitive conditions and at the same time so to arrange the systems that the same rates on traffic would provide substantially the same rate of return upon the values of the respective properties. Competition was therefore to be in service rather than in rates.

In accordance with this provision the commission adopted in 1921 a plan drawn up by Professor W. Z. Ripley of Harvard University as its tentative plan for the grouping of the railroads and in December, 1929, announced a definite plan of consolidation. It may be said that thus far the scheme of consolidation has been a complete failure. The success of the proposal depends upon the attitude of the railroads, for the commission has no authority to force consolidation, and the stronger railroads have not shown much appreciation of the advantages to be derived from absorbing some of the weak short-line roads at prices acceptable to the stockholders of the latter.

In fact the railroads have evinced more interest in plans of consolidation devised by and for themselves than in any comprehensive plan embracing the railroads of the entire country. With the assent of the commission a railroad may acquire control of one or more other roads through lease or stock control, and a number of roads have attempted to build up large systems by this means. Furthermore pooling arrangements are allowed with the approval of the commission, if they do not unduly restrict competition, and if a gain to the public can be shown Interlocking directorates are forbidden, but again broad latitude is given the commission in the administration of this provision.

Railway securities. The exclusive right of controlling new issues of rail securities was given to the commission. State control had proved itself in many cases to be arbitrary and discriminatory, and control by the federal authorities was adopted to ensure uniformity between the different railroads and to prevent the financial manipulation which was possible under the lax laws of some states. Power was given to determine whether the purpose to which it is proposed

to apply the proceeds of the issue is "reasonable, necessary, and appropriate" for the proper performance of public service, and if the application is approved the commission may make such terms or conditions governing the issue as it deems wise.

Railway service. Exceedingly broad powers with respect to car service were conferred on the commission. It may regulate the movement of cars and locomotives, compel the purchase of all classes of equipment, determine the compensation which one road shall pay for the use of the equipment of another, compel the common use of cars, locomotives, and terminals, and in cases of emergency issue priority orders or embargoes on traffic. The extension or abandonment of service facilities is likewise under its control; no new railroad may be built, no existing road may be extended, nor may any line be abandoned without the consent of the commission.

The railroad crisis. The transportation act of 1920 has not solved the difficulties of the railroads despite the constructive character of its provisions and the hopes of its sponsors. The roads as a whole were unable to earn the return expected even during the period of prosperity, and in the period of depression after 1929 the financial condition of many railroads became desperate, so desperate in fact that many were saved from bankruptcy and reorganization only by large loans from the Reconstruction Finance Corporation. In 1932 only a small minority of the Class 1 railroads [1] in the United States earned a net income, a large majority of them did not earn fixed charges, and a considerable minority even failed to cover operating expenses. The rate of return earned on property investment for the railroads as a whole, which had run between four and five per cent from 1926 to 1929, was down to one and one-quarter per cent in 1932. A number of railroads were in the hands of receivers, and the railroads had sought and secured 280 million dollars in short-term loans from the Reconstruction Finance Corporation.

The causes of the distress of the railroads then and today are many and can be mentioned only in general terms. One of the most

[1] Defined officially as including any railroad that has annual operating revenues of $1,000,000 or more.

general causes is lack of efficiency in management — the failure to effect available economies in operation and to modernize the services offered. This failure is ascribed by some to the paralyzing effects of government regulation. Critics also call attention to the alleged high salaries of railroad executives and to the wastes and costs of competition between railroads. There can be little doubt that the financial structure of some of the lines reflects unwise financial practices in the past and that the interest charges are beyond the normal capacity of the roads to meet.

More specifically the railroads have complained of new types of competition in transportation. Since 1920 there has been a considerable extension of pipe lines for the transmission of gas and oil, and the rapid development in the use of electric power has caused a decline in the tonnage of coal carried. Some lines have had to face what they consider unfair competition from river lines operating on inland waterways improved and maintained by large federal expenditures. Recent developments on the Ohio and Mississippi rivers have enabled river steamers to carry freight at rates which the railroads have claimed would be impossible without this indirect subsidy and which rail transportation cannot meet profitably.

But the type of competition which most seriously affects all railroads is that offered by the motor truck and the passenger bus, both in local and in long-distance services. Until 1936 the motor traffic, both freight and passenger, was under such regulation only as the individual states saw fit to apply to intrastate traffic, and it has moved on roads built by the states from general funds raised by taxation. Hindered by no federal requirements as to service or safety and subsidized to the extent of having its right of way built and maintained for it, the motor truck has offered a more flexible and cheaper service than the railroads and in many instances has been able to divert to itself the cream of the traffic.

The emergency railroad transportation act of 1933. The serious financial condition of the railroads as well as the obvious need of better coördination of the various means of transportation led Congress in 1933 to enact emergency legislation, intended to be in force for one year but extended by subsequent legislation to June 17, 1936.

The act provided for the appointment of a Federal Coördinator of Transportation and charged him with the immediate task of attempting to improve the condition of the railroads by devising measures of more economical operation. It was his duty also to attempt to reduce the fixed charges of the railroads by promoting the financial reorganization of their capital structures and to make studies of other methods of improving transportation as a basis for recommendations for legislation to be transmitted to Congress.

The staff of the Federal Coördinator was chiefly occupied with the investigation of the problem of waste in railroad operation and with recommending, for study and consideration by the railroads, means of achieving efficiency and economy by standardization of practices. Congress gave the Coördinator the power to issue orders to the railroads having the full authority of orders of the commission, unless revoked by the commission on appeal by the railroads, but it handicapped him considerably in his efforts to achieve economy by laying down the principle that economies should not be at the expense of labor. It was not permitted to reduce the number of laborers employed below that existing in May, 1933, except by the operation of natural forces such as death, retirement, or resignation, nor could a laborer's pay be cut below what he was receiving at that time. If in the interest of economy it became necessary to shift labor from one point to another, the financial burden was to be borne by the carrier concerned.

The office of the Federal Coördinator ceased to exist on June 16, 1936, and the railroads again were at liberty to effect economies by reducing the number or the pay of railroad laborers, but it is interesting to note that in the spring of 1936 an agreement was reached between the carriers and the representatives of labor whereby laborers adversely affected by such economies would be given appropriate compensation.

The emergency act amplified somewhat the powers of the Interstate Commerce Commission. The commission was given jurisdiction over railroad holding companies, with power to require that the stock holdings of such a company should be deposited with a trustee. Its authority to approve or to disapprove of mergers and consolidations was enlarged. It was instructed not to approve any

loan to a carrier by the Reconstruction Finance Corporation if in its opinion the public interest would better be served by the financial reorganization of the carrier.

Perhaps the most important provision affecting the commission was a change of the rule for rate-making laid down in the transportation act of 1920. This removed the requirement that a definite rate of return be allowed on the valuation of railroad property and required the commission to take into account, among other factors to which it should give due consideration in prescribing just and reasonable rates, the effect of the rates on the movement of traffic and the need of adequate service at the lowest cost consistent with the furnishing of such service by efficiently managed carriers.

The motor carriers act. The Federal Coördinator of Transportation recommended to Congress the enactment of legislation placing motor carriers and water carriers under federal control. Thus far Congress has not taken action on the bill for the regulation of water carriers, but the motor carriers act became a law on August 9, 1935. By this act all motor buses and trucks engaged in interstate commerce, with a few exceptions, as well as brokers who arrange for transportation but who do not themselves provide it, became subject to the regulation of the Interstate Commerce Commission.

The powers of the commission over motor carriers are similar to those which they have been exercising over railroads. To operate in interstate commerce a common carrier must secure a certificate of public convenience and necessity from the commission, while a contract carrier must secure a permit, and a broker a license. Rates charged must be just and reasonable and must not be discriminatory, and to enforce this provision the commission is given the power to prescribe both maximum and minimum rates. The commission may require financial reports and proof of financial stability; it may enforce uniform accounting methods; it may assume control of consolidations of companies and the issue of securities. It has the right to determine the qualifications to be demanded of operators, to specify maximum hours of work, and to prescribe such minimum standards of safety as it thinks the public interest requires.

The act does not contemplate exclusive federal control, for pro-

vision is made for the creation of joint state and federal boards, to be composed presumably of representatives of state commissions and of the Interstate Commerce Commission. Furthermore complete control of intrastate traffic is left in the hands of state commissions, although it is highly probable that this cannot remain unaffected by standards set up for interstate traffic.

The regulation thus provided does not touch upon one of the objections of the railroads to the motor traffic — the use of highways supported by general taxation — but this is scarcely within the province of the federal government. Nor is this act intended to bolster up rail finances by crushing the motor transportation facilities, for there is no intention of bringing the rates up to the level of the railroad rates. In determining rates which are just and reasonable the commission is to give due regard to the inherent advantages of motor transportation. It does however aim to achieve a greater degree of stability in transportation by eliminating the competition of the financially irresponsible, fly-by-night concerns and by forbidding disrupting price cutting, and it aims to make the highways safer by insisting on uniform standards of equipment and operation.

Labor legislation. The far-reaching effects on the economic welfare of the nation of a tie-up of railroad service resulting from labor disputes have impelled Congress from time to time to pass legislation setting up machinery for the settlement of labor disputes on the railroads or for improving the condition of the worker. It is interesting to note that as far back as 1888 Congress provided for voluntary arbitration in interstate railway labor disputes. From that time on various plans of conciliation, mediation, or arbitration have been put into effect. The most comprehensive of these earlier plans is found in the transportation act of 1920, which established a Railway Labor Board and authorized the formation of two types of boards of adjustment representative of the carriers and of the workers. Defects in the functioning of these organizations as well as a loss of confidence in the impartiality of the Railway Labor Board led to a new arrangement for settling railway disputes embodied in the railway labor act of 1926 — known also as the Watson-Parker act. This in turn has been revised and amended by legislation passed by Congress in 1934.

The Dill-Crosser railway labor act of 1934 creates a National Railway Adjustment Board with four divisions, each equally representative of the roads and the workers and having also a neutral member. In addition there may be established regional or system boards of adjustment if the railroads and the employees desire to set them up. If these voluntary boards or other agents of conciliation fail to secure a settlement of a dispute over the interpretation or application of existing agreements as to rates of pay, rules, or working conditions, the case goes to a division of the national Railway Adjustment Board for consideration and award. The board of mediation of five members, established by the act of 1926, is supplanted by a new National Mediation Board of three members to handle disputes involving changes in agreements as to rates of pay, rules, or working conditions.

This act marks also an important step in the process of unionization of railroad labor. It specifically prohibits any carrier from providing financial assistance from funds of the carrier to any union whatever. The carrier is also prohibited from interfering in any manner whatsoever with employees in joining or in refusing to join any organization or union, and the act further provides that the choice of representatives of any craft for the purpose of concluding a wage or other contract shall be determined by a majority of the employees voting. In other words the right of collective bargaining by majority representation is clearly guaranteed, and financial contribution by the company to a union of its employees is forbidden. The result has been a rapid movement in the direction of organizing national unions where these previously have not been found.

Old-age pensions. The same session of Congress enacted a law making retirement allowances for railroad employees mandatory. The act involved the pooling of compulsory contributions from the railroads and from the workers in a fund deposited in the United States Treasury. Retirement allowances were to be based both on the service period of the worker and on his average salary and were to be payable to any employee who had served thirty years or who had reached the age of sixty-five.

The railroads protested immediately that the burden imposed was too great for them to bear, and in May, 1935, the Supreme

Court declared the act unconstitutional on the grounds that the pooling provision involved deprivation of property without due process of law and that the purpose of the act lay outside the proper field of promoting or protecting interstate commerce. New legislation for the same purpose and with few changes was at once introduced into Congress and passed in August, 1935. This act having proved ineffective, the President called a conference of representatives of railroad management and railroad workers, which resulted in the formulation of the principles subsequently embodied in the retirement act of 1937.

The act of 1937 effected significant improvements in the administration of the railroad pension system. The pension program now covers not simply the employees of the railroads, but also those of auxiliary agencies, such as traffic associations and demurrage bureaus and national railway labor organizations. Annuities are payable by the United States Treasury to three classes of employees: those who on or after the date of enactment reach the age of 65, those who are totally and permanently disabled and have completed 30 years of service, and those who, being at least 60 years of age, have had 30 years of service or have become totally and permanently disabled, the amount of the annuity in this case being reduced in proportion as the applicant is under 65 years of age. The annuity for any individual is determined both by his years of service and by his average monthly compensation, with the proviso that no part of a monthly compensation in excess of $300 may be recognized for purposes of determining the amount of the annuity. The individual may choose, instead of a simple life annuity, an actuarially equivalent annuity payable to himself during life and to his wife after his death. The law also makes provision for the payment of death benefits.

The 1937 law provided for the transfer of individuals on the carriers' pension or gratuity rolls to the federal pension rolls, the pension payable in such circumstances not to exceed $120 a month; when pensions exceed that sum the excess must be paid by the carrier. In general it may be said that the pension is on a "pay-as-you-go" basis, with the income of each period intended to cover the costs of the system for that period. The 1937 act made however

a significant change by providing for the establishment of a Railroad Retirement Account in the Treasury for the purpose of accumulating a reserve to care for fluctuations in the charges that may accrue under the pension system.

In order to provide revenues to care for the obligations of the United States government under the Railroad Retirement Act, there was enacted as part of the pension program the carriers taxing act of 1937. This law provides for an income tax on employees and an excise tax on employers. The two taxes have the same base; *i.e.*, the compensation received by or paid to employees, and the same rate, which begins at $2\frac{3}{4}$ per cent of the monthly compensation not exceeding $300, remains at this figure through 1939, and thereafter rises gradually to $3\frac{3}{4}$ per cent after December, 1948. Both taxes are collected in the first instance from the employer.

As of November 30, 1937, the Railroad Retirement Board reported 73,378 cases on its books, of which 63 per cent were pensions, 36 per cent employee annuities, and 1 per cent survivor or death benefit annuities. The aggregate of monthly annuities and pensions was then $4,472,588.09. As an indication of the amount of the individual payments, the board reported as of October 31, 1937, that 4 per cent of the pensioners were receiving less than $20, 29.9 per cent from $20 to $39.99, 34.9 per cent from $40 to $69.99, and 5.7 per cent $120. Thus far the average age of retirement has been 66.7 years.

A new crisis in transportation. Following the upswing in business after 1933 the acute financial distress of the steam railroads was somewhat alleviated, but with the downturn in general business which began in 1937 an even more serious financial crisis developed. Of the 782 operating railroad companies in the country, only 139 are Class I roads, but these companies operate 93 per cent of the country's railroad mileage. These Class I railroads sustained a deficit of 139 millions in 1932, then earned an aggregate net income of 164 million dollars in 1936, but their net income declined to 98 millions in 1937. In that year 69 of the Class I railroads reported total net income above fixed charges of nearly 231 millions, but 62 systems reported an aggregate deficit of over 132 millions. Shrinking revenues and rising costs accentuated the downward trend, so

that in the first three months of 1938 the Class I railroads as a whole failed to meet fixed charges, accumulating a deficit of over 100 million dollars. In March, 1938, a special committee of the Interstate Commerce Commission reported that 37 Class I railways were in the hands of the courts. The bonded indebtedness of these roads was approximately 3.8 billion dollars, the accumulated unpaid interest, 600 millions, and the excess of interest coming due over interest payments some 100 millions yearly. Later statistics from the Bureau of Information of the Eastern Railways reported that of the total railroad mileage, 30.67 per cent was operated by receivers or trustees as of June 1, 1938; comparative percentages for December 31, 1932, were 8.71, and for December 31, 1929, 2.19.

The immediate and proximate cause of the crisis is to be found in the decline in railway traffic after March of 1937, a decline which became precipitate in the last two months of that year. Thus the freight revenues of the railroads for February, 1938, were more than 25 per cent below those of February, 1937, and the total passenger revenues declined nearly 5 per cent for the same period. The financial strain was enhanced by increases in expenses for both labor and materials. The effect of the decline in revenues and the increase in expenses was reflected in the diminished net railway operating income for January, 1938, — $6,920,000 compared with $38,867,000 for January, 1937. The New England region and the entire western district did not have sufficient revenues to cover their operating expenses and taxes in January, leaving nothing available for interest charges.

One effect of these curtailments in railway income has been reflected in the decline in railway employment. Between February, 1937, and February, 1938, there was a reduction of 14.3 per cent in the total number of employees. The railroads have found it necessary also to curtail substantially their expenditures for maintenance of ways and structures and for the replacement of equipment. Although existing equipment is adequate to handle the present volume of traffic, a revival of business activity would find the roads suffering from a shortage of equipment. Further it is authoritatively stated that large investments are needed for modernizing equipment and shops, not simply to place the roads in position to

handle more traffic, but to enable them to handle present traffic at lower cost and with greater expedition.

A program of relief. At the request of President Roosevelt a special committee of three members of the Interstate Commerce Commission was established to advise on measures of immediate relief for the railroads. A five-point program was offered in 1938:

1. The improvement of railroad equipment and facilities was advocated as an urgent necessity, both to enable the railroads to maintain their competitive position and to enable the carriers to render more satisfactory service, and it was suggested that the purchase of shop equipment and appliances and cars and locomotives could be financed by government funds in the amount of $300,000,000, with the equipment itself pledged as security for such advances.

2. Under existing law loans can be made by the Reconstruction Finance Corporation only with the approval of the Interstate Commerce Commission. Furthermore, with the exception of loans to receivers or trustees and loans for the purchase and maintenance of equipment, the commission must certify that the railroad may reasonably be expected to meet its fixed charges without reorganization. It was suggested that the Reconstruction Finance Corporation might safely make short-term loans in cases where the prospective earnings of the applicant, together with the character and value of the security, furnish reasonable assurance of the applicant's ability to repay the loan within the time fixed To afford added protection for such government loans the obligations to the RFC might be given a position of priority in bankruptcy proceedings.

3. The effect upon general credit conditions of existing and threatened defaults on interest bearing obligations was recognized as one of the most important aspects of the railway problem Several suggestions were advanced with the objective of improving railroad credit and thus bolstering general credit conditions. Two of these received serious consideration: (1) a government subsidy to the railroads, based in individual cases on the value of the property, earnings for a test period, or fixed charges, and (2) the use of government credit to underwrite or guarantee bonds issued in voluntary reorganizations to reduce fixed charges.

4. For many years government traffic has received substantial reductions in rates when moving over sections of railroads built with the aid of land-grants from the government, competing carriers making similar rate concessions in order to share in government shipments. In 1937 these reductions amounted to $7,000,000. As an immediate aid to railroad revenues it has been suggested that for the future the government pay the same rates as other shippers.

5. The reorganization of the railroads in receivership has become a pressing public problem. While recognizing that reorganizations must be slow pending the stabilization of conditions which would make possible a forecast of future earnings, the committee of the Interstate Commerce Commission recommended certain improvements in the federal reorganization laws together with the establishment of a single court to have jurisdiction over railroad reorganizations.

Though the railroads have sought and obtained certain rate increases to improve their inadequate earnings, obviously rate increases, save as applied to carefully selected non-competitive traffic, offer no solution to the carriers' financial problem. Some indirect relief however has come from the increases in rates prescribed for competing transportation agencies.

Long-term program of railroad rehabilitation. The permanent rehabilitation of the railroads can be assured only by the consideration of the "railroad problem" as part of a national transportation problem. Such a permanent solution of the transportation problem must, according to the views of the special committee of the Interstate Commerce Commission, proceed along the following lines: establishment of an increasing measure of coördination among the different competing transportation agencies; improvement in service and reduction in the cost of railway operations through coördination, consolidation, or economies effected by the railroads themselves; a recasting of the financial structures of the railroads to effect reductions in their fixed charges; extension of the jurisdiction of the Interstate Commerce Commission over water carriers, air carriers, companies which supply services ancillary to transportation, and all non-carrier subsidiary companies of carriers, so as to provide for the establishment of uniform regulation of

all transportation instrumentalities; and the elimination of such financial abuses as still exist.

Conclusions. The policy of the federal government has changed in the course of sixty years from one of indifference to one of minute control of most phases of railroad management. While early regulation stressed the necessity of maintaining competition between railroads, later policy has recognized the futility of trying to maintain competition in rates and has emphasized competition in service. Regulation has not as yet brought about a satisfactory solution of the railroad problem, but it must be recognized that the modern railroad problem is essentially different from that confronting the country when the interstate commerce act was passed. The abuses of the earlier period have in the main been dealt with successfully; today the problem is the important and difficult one of giving the railroads financial stability.

The public has a vital interest in a successful solution of this problem. As of December 31, 1936, the total railway capitalization was $21,961,035,000, of which about 12 billions was funded debt and nearly 10 billions stock. The annual reports of the railroads showed that there were 855,261 stockholders in 1936 and that approximately 65 per cent — $6,617,000,000 — of the railway bonds were held by insurance companies, banks, endowed educational institutions, and foundations. The holdings of life-insurance companies alone amounted to $3,267,000,000. As investors therefore the public is immediately concerned with the financial problem of the railroads.

As shippers the public is interested in the maintenance of an adequate, efficient, and economical transportation system. Inasmuch as the railroads now are, and so far as one can see will continue to be, the backbone of the transportation system, anything which threatens their continuing effectiveness is a matter of grave concern. The railroads are also large purchasers of equipment, particularly from the heavy industries, and large employers of labor, so that any decline of railway purchases in the goods and labor markets must have serious repercussions on related industries.

It is in consideration of the far-reaching consequences of distressed conditions in the railroad industry that agitation for gov-

ernment ownership and operation of railroads secures much of its support. Despite the contention of some that the present plight of the railroads is the result of so much regulation that initiative in management has been destroyed, unregulated railroad operation is a practical impossibility in the United States. Politically the only alternative to continued regulation is government ownership and operation. Some time ago the Federal Coordinator expressed his conviction that this is an inevitable step in the future. Railroad executives and many others disagree emphatically. They fear that government operation would be even more rigid and unprogressive than has been private operation under government regulation; they fear the political power which such a step would give to railway labor, as well as the precedent which might be set up for still further extensions in the scope of government activity. Nevertheless it must be recognized that continued failure to find a solution to the present problems of the railroads is likely to keep alive and strengthen the demand for government ownership and operation.

XLV

THE AMERICAN FARM PROBLEM

In agriculture we find none of the evidences of monopoly which have been so disturbing in industry. Agriculture's problems arise rather from the failure of competitive price to bring about easily and quickly the adjustments in production which would preserve a normal balance between costs and prices. In this chapter we are concerned first with the nature of the special economic problems of agriculture and secondly with recent attempts to solve these problems by governmental action. Limitations of space will make it necessary to confine our discussion chiefly to recent events in American agriculture and to certain only of the major crops, especially wheat and cotton.

Agricultural maladjustment: Basic causes. We should perhaps note first of all that in the case of agricultural products normal equilibrium is reached with more difficulty than in the case of industrial products, because of the vagaries of nature's bounty, the inelasticity of demand which frequently prevails, the highly individualized control of productive efforts, the less compelling influence of marginal costs on productive activity, and the greater inflexibility in the transfer of productive agents from use to use. Increases in the prices of agricultural products do indeed produce about as ready a response as the growing seasons and the weather permit, but decreases in price, even when so drastic as to wipe out the margin of profit of many producers, will not cause an immediate readjustment of supply to the new conditions.

These facts are illustrated in the history of American agriculture. Taking wheat as an example, we find that from 1900 to 1914 there was a slight decline in the area planted to wheat, with total production remaining roughly stationary, and exports forming a decreasing percentage of total production. The World War was accompanied by reduction of wheat culture by the belligerent nations and the removal of several important suppliers from the world market.

The demand of the warring countries was therefore in large measure focused on the United States and reflected in a higher price for wheat. This was of course a period when the general level of prices was rising, but the wheat increase outstripped that of commodities in general; whereas the index of wholesale prices shows an increase of 118 per cent from 1914 to 1919, the Chicago basic price of wheat rose from an average of $0.988 a bushel in 1914 to $2.32 in 1919, an increase of 134 per cent. The acreage of wheat harvested in the United States increased from an average of 48 million acres in the period 1909–1913 to an all-time high of over 77 million acres in 1919, and production, which on the average was under 700 million bushels in the pre-war period, rose to an average of around 800 million bushels in the war and immediate post-war period, with record yields in 1915 of over a billion bushels and in 1919 of 952 million bushels.

From 1920 to 1921 the price of wheat decreased from an average of $2.43 a bushel to $1.36 and continued to decline until 1924. In 1925 there was a material increase, but in 1926 the price again began to fall, reaching its low point of $0.524 in 1932. The relationship between prices, acreage, and production is brought out in the table on the opposite page.

The price decline is explained by a number of factors. Chiefly it is the result of changes in the value of the dollar. Except for the upward jump from 1923 to 1925 and the compensating gradual decline since 1925, the movement of wheat prices corresponds closely to the changes in the general price level. As to factors peculiarly affecting wheat, we note that European countries returned to production after the war and that an increasing number of these nations stimulated the domestic production of wheat and of other foodstuffs by increasing tariff barriers. Canada, Australia, and Argentina, producing with lower costs than in general obtain in the United States, competed vigorously for what was left of the European market, and eventually Russia and the Danubian countries began to be felt again in world markets. In the United States changes in dietary habits caused a decrease in the per capita consumption of wheat, while in many European countries disordered internal conditions prevented an expansion in wheat consumption

which might otherwise have been expected. The net result of a marked increase in world production of wheat, not balanced by an equivalent increase in demand, was the piling up of surplus stocks.

WHEAT PRODUCTION IN THE UNITED STATES [1]

Year	*Index of wholesale prices 1926 base*	*Average price of wheat (Chicago basic price, in cents per bushel)*	*Area planted (in thousands of acres)*	*Production (in thousands of bushels)*
1920	154 4	243 8	67,977	843,277
1921	97.6	136 2	67,681	818,964
1923	100.6	111 1	64,510	759,482
1925	103 5	163 0	61,736	669,142
1927	95 4	135.0	65,649	874,733
1929	95.3	121 2	66,787	822,180
1931	73 0	66.5	65,618	932,221
1932	64.8	52 4	64,927	745,788

Maladjustment in wheat growing after 1920. Reference to the table above will show quite clearly that American wheat growers failed to make adequate adjustment to changing prices. All through this period production of wheat stayed about thirty-five per cent above the average of 1909–1913. The especially tardy adjustment of agricultural production may be ascribed to two important causes. The first is that leaders of agriculture seem consistently to have misread the situation. In 1920 for example editors of farm papers and other leaders appeared to have no misgivings regarding the export market for farm products, and throughout the following decade the feeling among both the leaders in farm affairs and the professional speculators was that the export market was bound to recover. "The prevailing opinion, in both trade and government circles, was that wheat prices were abnormally low in 1928–29; but in retrospect it is clear that speculative forces held them above levels appropriate to the real conditions. The danger signals given by unprecedentedly high 'visible supplies' of wheat were largely ignored . . . the conviction prevailed that 1929–30 would see higher prices and ready absorption of surplus wheat stocks. This was the situation when the Federal Farm Board came

[1] Derived from tables given in J. S. Davis, *Wheat and the AAA*, Brookings Institution, Washington, 1935, to which we are indebted for much information regarding the wheat program. This is the first of a series of studies of the operations of the AAA in controlling production.

into existence in July, 1929, and when the financial crisis broke in the New York stock market in October [of that year]. In his annual report for 1929, written shortly after this, the Secretary of Agriculture observed that the agricultural situation was remarkably free from evidence of overproduction."[1]

The second reason which may be given to explain the continued maladjustment in agriculture is the conviction throughout the period that prices were unjustly low and the expectation that Congress would do something for agriculture. Constant pressure by farm groups was exerted, and a number of measures supposed to aid agriculture were either enacted into law or passed by Congress, only to be vetoed by the President. The first of these measures to be enacted was the emergency tariff act of May 27, 1921, which placed a duty of 35 cents a bushel, later raised by President Coolidge to 42 cents, on the import of wheat. The inevitable failure of this action to improve the situation of the wheat farmers led to the passage of the two McNary-Haugen bills designed to raise the domestic price by subsidizing the export of wheat. These bills were vetoed by President Coolidge in 1927 and 1928. Finally under President Hoover the agricultural marketing act of 1929 was passed, and the active participation of the government in stabilization projects began. It is not our purpose here either to describe these different projects or to pass judgment on the plea of agriculture for special treatment at the hands of the government but merely to record the fact that a decade of political agitation and political promises could not fail to influence the policy and behavior of the farmers.

Even without the depressing effects of the world crisis American agriculture would have had difficulty in solving the problems that confronted it in the years following 1929. The world crisis however put an end to American loans to foreign countries, causing collapse of the export market for American produce, while internal disturbances wrought havoc with the domestic market. Prices in general declined, but farm prices fell even more rapidly, so that there were drastic reductions not only in the size of the farmers' income but in its purchasing power as well.

[1] J. S. Davis, *Wheat and the AAA*, 1935, p. 17.

Farm debt. The farmers' situation was aggravated by an unusually heavy burden of debt, estimated at a total of twelve billion dollars in 1933. Part of this consisted of short-time loans, but the major portion was in the form of farm mortgages, which were to a considerable extent, as a glance at the following table will indicate, the result of purchases of farm lands at the speculative prices obtaining during the war period.

TOTAL FARM MORTGAGE DEBT

Year	*Millions*
1910	$3,320
1920	7,859
1925	9,361
1928	9,468
1930	9,214

The interest on this debt amounted in 1929 to about half a billion dollars yearly. With farm income curtailed, many farmers found themselves unable to meet their interest payments, and foreclosures followed Between 1920 and 1930 nearly half a million farmers lost their farms through foreclosure and forced sale, and many formerly independent farmers were forced into the class of tenants. The extent of this is seen in the fact that in 1910 nearly two-thirds of all farmers were owners, whereas in 1932 over half were tenant farmers; the social loss from a movement of this character is obvious. Between 1930 and 1932 the farm mortgage debt decreased to around eight and one-half billion dollars through continuation of the process of foreclosure and sale, but even so not all of the delinquent mortgages were foreclosed. Financial agencies reporting in 1933 to the Department of Agriculture showed that 36 per cent of all such loans were delinquent, with the figure running as high as 55 per cent in the wheat area of the Dakotas Mortgage moratoria, some of them later invalidated by the courts, were established by state law in many of the agricultural states, and in their absence organized mass movements by farmers interfered with court proceedings and halted foreclosures and forced sales.

Farm income and prices. A gross farm income, estimated at twelve billion dollars in 1929, fell to eleven billions in 1930, to

eight and one-half billions in 1931, and to six and one-half billions in 1932, a decline of 46 per cent in the four-year period. In comparison it is to be noted that estimates of the Department of Commerce and of the National Bureau of Economic Research put the total income of American enterprise at eighty-three billion dollars in 1929, with an aggregate reduction of 40 per cent by 1932. In the fall of 1933 general commodity prices were 30 per cent under their 1926 level, farm products 46 per cent, building materials 16 per cent, metals and metal products 18 per cent, and textile products 24 per cent. But these figures do not quite accurately portray the seriousness of the changes in the position of the farmer, because the prices of farm products had already fallen much farther from the peak of 1920 than had the prices of industrial products. This relationship is brought out more clearly in the following table, which takes the pre-war period of 1909 to 1914 as a standard for comparison.[1]

INDEX NUMBERS OF

Year	*Prices received by farmers*	*Prices paid by farmers*	*Ratio of prices received to prices paid*
1909-1914	100	100	100
1920	211	201	105
1925	156	157	99
1928	149	155	96
1930	126	145	87
1932	65	107	61

The Federal Farm Board, 1929–1933. The first broad act for the relief of agriculture was the agricultural marketing act passed in the spring of 1929. Beginning in 1930 the grain and cotton stabilization corporations of the Federal Farm Board, created by that act, did considerable buying of wheat and cotton. From July, 1930, to August, 1932, over one million bales of cotton so bought were being held. The bulk of this was distributed, after processing, through the National Red Cross, and the remainder, together with one and one-half million bales which were held as a result of collateral loans, was turned over to the Department of Agriculture in 1933 for the cotton

[1] United States Department of Agriculture, *The Agricultural Situation*, June, 1938, p. 24.

option pool under the new farm legislation. The Department of Agriculture itself had accumulated since 1930 over half a million bales, taken in lieu of cash on repayments of collateral seed loans to cotton growers (such loans were being made in October, 1932, on a nine cent per pound basis for cotton with the market price at six cents), and this stock was also turned into the option pool. There were other stocks still hanging over the market and available for the pool, having been held by coöperative associations on loans granted by the RFC. As to wheat, the accumulated holdings of the Farm Board were liquidated in 1932 and 1933, largely by distribution through the National Red Cross after conversion into flour. By the end of April, 1933, the Grain Stabilization Corporation had closed out all its holdings, both spot and future. The Federal Farm Board came to an end in May, with its usable credit functions salvaged for the new Farm Credit Administration. The heavy surpluses of wheat and cotton which had been kept hanging over the market and the failure to offer any discouragement to increases of production had helped to create an almost continual decline in the prices of these commodities during the life of the Board, and the Board's operations did not tend to decrease wheat production. The final loss on the 500 million dollar revolving fund for price stabilization was about 350 million dollars.

The farm program of the Roosevelt Administration: General objectives. The objective of the government's farm program in 1933 can be simply stated. It was to bring, with as little delay as possible, the prices of farm products into a more favorable balance with the prices of other goods and to stabilize this more favorable relationship once it had been attained. The agricultural crisis was deemed so serious that various means of achieving this objective were tried simultaneously. In so far as the program for raising the general price was based on the assumption that, once prices started moving upward, the more depressed prices would make the greatest relative recovery, the general price program should be considered a part of the farm program. New agricultural credit agencies, liberally supplied with government funds and under government control, were established to make loans to the farmer and thus to give him sufficient financial strength to hold his produce for a

favorable market. Devices for enlarging foreign markets, such as the negotiation of reciprocal tariff agreements authorized by the tariff act of 1934 or the export of agricultural produce with government subsidy, were also directed at this same end.

The main program however sought (1) to make it immediately profitable for producers of the basic agricultural commodities to reduce production voluntarily, (2) to check the development of unfavorable competitive conditions, and (3) to supervise the growth of equitable distributive processes for all agricultural commodities. This program was embodied in the agricultural adjustment act of 1933 and its various amendments.

The agricultural adjustment program of 1933–1936: Legislation: The act of 1933. The agricultural adjustment act of 1933 was signed by the President on May 12, 1933; it was amended and supplemented by further legislation in 1934 and 1935. On January 6, 1936, the most significant provisions of the act of 1933 were declared unconstitutional by the Supreme Court. But for nearly three years the administration had been able, by virtue of the authority conferred by this act, to exercise a high degree of control over American agriculture. Enduring changes in the pattern of American agriculture and in the point of view of the American farmer could not fail to result. Furthermore the economic philosophy which inspired the drafting of the act of 1933 did not die with the court decision. It persisted and found tangible expression in the subsequent farm legislation of 1936 and 1938. A critical study of the act of 1933 and supplementary legislation up to the end of the year 1935 is therefore of more than historical interest. It provides an almost essential background for an understanding of the objectives of the federal government and for any evaluation of the probable results of the present laws affecting agricultural production.

"Parity prices." The agricultural adjustment act was entitled "An act to relieve the existing national economic emergency by increasing agricultural purchasing power," etc. Its major goal was restoration of the purchasing power of basic farm products to the level of the pre-war period, August, 1909, to July, 1914, except in the case of tobacco and potatoes, for which the base period was from August, 1919, to July, 1929. Originally the list of basic farm

products comprised wheat, cotton, hogs, field corn, rice, tobacco, and milk and its products, but it was extended by later legislation to include beef and dairy cattle, rye, barley, flax, grain sorghums, peanuts, sugar and sugar cane, and potatoes.

For each basic commodity there was at any given time a so-called "parity price." For example in the base period the average price of wheat was 88.4 cents a bushel. In order that a bushel of wheat should have as great a command over commodities in general as it had in the base period, the price in April, 1936, for example, would have to be $1.10½. In this same month the parity price for cotton was 15.5 cents a pound, compared with 12.4 in the base period, and so on for the other basic commodities. It is to be noted that the plan did not attempt to set fixed prices but sought to adjust farm prices to the general price level, whatever that might be. Thus by January, 1937, the rise in the general price level had brought the parity price of cotton to 16.4 cents a pound.

Voluntary production control. In order to limit production the Secretary of Agriculture was empowered to provide, by voluntary agreements with producers, for the reduction of acreage or production for market of the basic commodities and to compensate the coöperating producers by the payment of rent, if the land was leased out of production, or by benefit payments on the restricted allotments grown. In the case of cotton the Secretary was empowered to use option contracts, alone or in conjunction with rental and benefit payments. Under these contracts the cotton grower who agreed to a specified reduction in his acreage planted to cotton, with no increase in the use of commercial fertilizer per acre, received a call on an equivalent amount of cotton held by the government. Thus if the price of cotton should rise and the grower exercise his option, he would profit by the amount of the price increase, whereas he was not held liable for any loss in the holding of cotton if the price declined. Under this option plan the grower was required also to agree that he would not use land taken out of cotton production to grow cash crops or any other nationally produced agricultural commodity. It is interesting to note that in the 1933 reduction of cotton production, sixty per cent of the growers chose a combined cash-and-option plan, while the rest took full cash payments.

Compulsory restriction of production. A significant departure from the method of voluntary crop reduction was embodied in several later measures. The Bankhead cotton control act of April, 1934, applicable originally to the crop years of 1934 and 1935 but extended for another year by legislation in 1935, definitely set 10 million bales as the allowable crop from the 1934 planting and imposed a ginning tax of 50 per cent of market value, or a minimum tax of 5 cents a pound, on all cotton ginned in excess of this amount. Administration of the act involved the determination of a quota of bales which might be marketed tax-free for each state, with subquotas for counties and individuals. These quotas were based on average production from 1928 to 1932.

Compulsory production control was made applicable also to sugar by the Jones-Costigan sugar act of May, 1934, to tobacco by the Smith-Kerr tobacco control act of 1934, and to potatoes by a 1935 amendment to the agricultural adjustment act.

Supplementary powers of control. The Secretary of Agriculture was permitted also to make marketing agreements, exempt from the operation of the anti-trust laws, with processors and handlers of any agricultural commodity or its derived products "in the current of interstate or foreign commerce." The 1935 amendment to the agricultural adjustment act strengthened the power of the AAA to deal with the producers of milk, fruit, vegetables, soy beans, pecans and walnuts, tobacco, and naval stores by this method.

By the original act of 1933 the Secretary of Agriculture was empowered to use the proceeds of all taxes authorized by the act for the purpose of expanding the foreign market for American produce and thereby removing the surplus from the domestic market. It was by virtue of this authority that an arrangement was made in 1933 for the export of surplus wheat from the Pacific northwest by the North Pacific Emergency Export Association. An average loss of 22 cents a bushel was made good out of funds derived from the processing tax described below. An extension of this power to develop foreign markets is found in the amendment of 1935 which permitted the use of thirty per cent of the receipts from import duties to subsidize the export of farm products, with

the exception of raw cotton, or to increase the returns of producers on the part of the crop consumed in this country. This amendment is still in force, although it has been modified to permit the subsidized export of raw cotton.

Another provision of this amendment strengthening the control over domestic prices was the authority granted to set import quotas for goods if they should enter in such quantities as to threaten the price structure of American-grown products.

Processing taxes. The cost of the payments mentioned above was to be met from taxes on the first domestic processing of the products. The Secretary of Agriculture was given the power to set the amount of the tax on each basic commodity at any figure up to the amount of the difference between the current average farm price of the commodity so processed and the parity price. The processing taxes were remitted on products exported and were not levied on supplies processed by the grower for his own consumption or used for charity. A compensating tax was levied on any imported manufactured article whose value was chiefly derived from one of the commodities concerned. In addition to these provisions, compensating taxes were to be used as a safeguard against the defeat of the purposes of these measures by excessive shifts in consumption to non-taxed competing commodities.

Operation and results: The International Wheat Conference. Up to this point we have been passing in review the principal features of the legislation of 1933–1935. We must now give attention to the operation of these laws and their results. The program for the reduction of wheat production as first applied was an outgrowth of an international conference on wheat control. Discussions on this topic were initiated in the summer of 1933 in connection with the World Monetary and Economic Conference, and in September, 1933, an agreement was finally signed by twenty-one nations calling for the reduction of production by Canada and the United States to the extent of fifteen per cent of average production from 1931 to 1933, with a reduction of exports by other producers, and with a provision for the lowering of tariff duties by importing countries after the price had advanced to around sixty cents a bushel, a figure about fifty per cent above the then current market price. The allot-

ment of export quotas remained unsettled pending further discussions, but provisional quotas were assigned, that of the United States being 47,000,000 bushels. Agreement on these quotas was however never reached. Argentina asked for a larger allotment than Australia, Canada, and the United States were willing to concede and actually exported more than the provisional quota permitted. In December, 1934, the attempt at international control collapsed: Canada and Australia both expressed dissatisfaction with their quotas, and Argentina definitely refused to agree to a continuance of the plan.

The wheat program: Operation. The international agreement coincided with the development of the AAA program. The Secretary of Agriculture had already called upon the wheat farmers to sign contracts calling for a reduction in the area to be sown to wheat in 1934 and 1935 of not more than twenty per cent of the average area sown in the three preceding years. Compliance was to be rewarded by the payment of benefits by the AAA. With the conclusion of the international agreement a fifteen per cent reduction was called for; in other words, the contracting farmers agreed to sow to wheat eighty-five per cent of the average area which had been devoted to wheat in the period 1930 to 1932.

The payment of benefits to contracting farmers was based on what was called the domestic allotment plan. For the country as a whole the domestic allotment was that portion of the total wheat production which entered into domestic consumption. The period 1928–1932 was taken as the basis of the computations, and it was estimated that fifty-four per cent of the wheat produced was consumed at home. State allotments, county allotments, and individual allotments then were made. For an individual farmer the allotment was a rough approximation of fifty-four per cent of his own production in that base period. Benefits were paid, not on the total production of a farmer under contract with the AAA, but on his allotment. For a part of his production therefore he received the open-market price, whatever that might be, but for the remainder he got the open-market price plus a benefit. To secure funds to pay this benefit a processing tax of thirty cents a bushel was levied on the first processor; of this two cents was deducted for the

promotion of exports, leaving a benefit payment of twenty-eight cents a bushel less a further deduction for local administrative expenses. In May, 1934, the benefit was increased to twenty-nine cents a bushel. The effect of this was intended to bring the income of coöperating farmers up to the "parity" level on their shares of wheat consumed domestically for food.

For the 1935 crop year a ten per cent reduction of the area sown to wheat in the base period was required, and it was announced that the benefit payments would amount to thirty-three cents a bushel, without any change in the processing tax, giving a total benefit of over $115,000,000. During the summer of 1935 it was announced that contracts covering the period from 1936 to 1939 would be offered the farmers calling for a five per cent reduction in area sown to wheat in 1936 and permitting the administration to require a reduction in subsequent years as high as twenty-five per cent in place of the maximum of twenty per cent embodied in the contracts signed in 1933. Dissatisfied farmers would be permitted to withdraw at the end of the first two-year period. The signing of contracts under this program was nearly completed when, on January 6, 1936, this form of production control was declared unconstitutional. Subsequently authority was granted by Congress whereby those signers who had observed these provisions in their fall sowings prior to January 6, 1936, could receive 21.5 cents a bushel on certificates of compliance.[1]

Results of the wheat program. The wheat program inevitably affected commodities other than wheat and persons other than the producers or handlers of wheat. Its influence might therefore be spread over a very wide area and be felt in many subtle ways in other fields. At this point we are not interested in an appraisal of the economic bearings of the agricultural program as a whole, of which the wheat program was perhaps the most important part, but in discovering how successful the agricultural administration was in the light of its own short-run objective of adjusting the production of wheat by means of the payment of benefits so that the effective return per bushel of wheat would have, for the con-

[1] E. G. Nourse, J. S. Davis, and J. D. Black. *Three Years of Agricultural Adjustment Administration*, Brookings Institution, 1937, p. 94.

tracting farmers, a purchasing power over other commodities more nearly in accord with that obtaining in the base year.

So far as the reduction of acres sown to wheat was concerned the program was not successful. The 1934 program called for a 15 per cent decrease in acres planted to wheat by farmers signing contracts with the government. Despite the fact that farmers controlling 75 per cent of the wheat acreage of the base years, 1928–1932, signed these contracts, increased plantings by non-signers resulted in a net decrease in 1934 of 4 per cent only. The 1935 program called for a reduction in acreage of 10 per cent by signers, whereas actually the acreage planted was in excess of that of the base years. Production however did decline from an average of 864,000,000 bushels in the base period to an estimated yield of 496,000,000 bushels in 1934 and 626,000,000 in 1935 – crops small enough to cause a rapid decrease in the large carry-over of wheat in 1933. This might be considered evidence of success of the program if one failed to remember that unfavorable weather conditions prevailed during these years. The investigation of the Brookings Institution indicates that had there been no attempt to control production the 1934 crop might have been 25 or 30 million bushels and the 1935 crop perhaps 20 million bushels larger than they actually were In other words the effect of the AAA measures of restriction were practically negligible in comparison with the effects of weather conditions.[1] There is no way of appraising accurately what the AAA program might have accomplished had favorable weather conditions prevailed, but the Brookings Institution investigators believe it quite improbable that any voluntary program of restriction could have succeeded in reducing production by an amount sufficient to relieve the government of the embarrassing problem of disposing of a large surplus of wheat over and above domestic requirements.

The cotton program: Operation. The special difficulties of the cotton planters (in addition to those arising from conditions affecting industry generally) may be traced in part to the influence on acreage and production of high war and post-war prices, in part to the increasing use of substitutes for cotton, and in part to the in-

[1] E. G. Nourse, J. S. Davis, and J. D. Black, *Three Years of Agricultural Adjustment Administration*, 1937, p. 127.

crease in the culture of cotton in other parts of the world. Even before the depression the plight of the cotton planter was attracting national attention, and under the agricultural marketing act of 1929 a Cotton Stabilization Corporation was set up to buy cotton and hold it for more favorable prices. The depression brought a reduction in the consumption of cotton both at home and abroad, a sharp drop in the price of cotton, and a decrease in the incomes of cotton planters. The price of cotton decreased from an average of 18 cents a pound in the season 1928–29 to a record low of 4.6 cents in June, 1932, and the planters' income from cotton declined from $1,511,950,000 in 1929 to $460,694,000 in 1932. Production however continued without much curtailment, with the result that the world carry-over of American cotton amounted in 1932 to nearly 13,000,000 bales as contrasted with a carry-over of 5,114,000 in 1928. Nevertheless, encouraged somewhat by an improvement in prices and the hope of a general recovery in business conditions, cotton planters had in 1933 put 41,000,000 acres into cotton; this would have given an estimated yield of seventeen and one-half million bales.

It was to meet this critical situation that the Agricultural Adjustment Administration formulated its plan for an immediate reduction of cotton planting by offering a cash or a cash-and-option payment to planters who would agree to plow under not less than 25 per cent nor more than 50 per cent of the growing cotton. Ten and one-half million acres of cotton were thus taken out of cultivation during the summer of 1933, but even so the crop amounted to 13,-177,000 bales. Early in 1934 a million growers signed contracts with the AAA which required a reduction in acreage in that year by 40 per cent of the average acreage under cultivation in the base years, 1928-32, and permitted a requirement of a 25 per cent reduction for the base acreage in 1935. These contracts covered about 38 million acres out of the average total of 41,437,000 acres devoted to cotton in the base period. Unhappily for the success of this plan of voluntary adjustment, non-contracting growers materially increased their plantings in 1934. It was therefore felt necessary to enact the Bankhead cotton control act in April, 1934, setting a definite tax-free quota of 10,000,000 bales of 500 pounds, and

providing for the subdivision of this quota into state, county, and individual allotments. The Bankhead act was not a substitute for the subsidized reduction in the area planted in cotton, but supplementary to it and intended to control the production of non-coöperators.

The cotton program was supplemented by generous loans on farm-stored cotton made by the Commodity Credit Corporation. Loans "without recourse" to the planter were offered in the fall of 1933 on a basis of 10 cents a pound; the next year the basis was changed to 12 cents a pound. "Without recourse" means that if at the maturity of such a loan the market price of cotton should be lower than the basic price, the borrowing farmer could completely settle his obligation to the government by simply forfeiting the cotton pledged as security. Since the market price of cotton was below this pegged price, planters eligible for loans naturally turned to the government. The result was that the government was called upon to make the loans on roughly half of the 1934 crop.

A variation of this plan was announced for the ensuing year (1935-1936). The Commodity Credit Corporation was to lend on the basis of 10 cents a pound to cooperating producers, but in order to assure producers a minimum price of 12 cents a pound the AAA was to grant them, after the sale of their cotton, the difference between the market rate and the minimum price, up to a maximum grant of 2 cents a pound. The effect of this was to peg the price at 10 cents a pound in the open market. The price might of course rise above this, but if it were to fall cooperating producers would find it more profitable to secure a loan from the Commodity Credit Corporation and wait for better prices or, failing these, allow the government to take possession of the cotton.

Appraisal of the cotton program. There can be no question of the accomplishment of the AAA in effecting a marked reduction both in the area planted to cotton and in the quantity of cotton produced. The results of contracts with the government to destroy crops in 1933 and to rent acreage to the government in 1934 and 1935 are seen in the following table, in thousands of harvested acres and thousands of bales of 478 pounds net: [1]

[1] Nourse, Davis, and Black, *Three Years of the AAA*, 1937, p. 138.

Year	*Acreage*	*Production*	*Tax-exempt quotas*
1930–32 (average)	39,013	14,677	
1933	29,383	13,049	
1934	26,866	9,636	10,460
1935	27,335	10,638	10,983

It is estimated that in the absence of this control production might have been from 10 to 13 million bales greater in the three-year period 1933–35 than it actually was.

The program also added substantially to the income of the cotton planters. The curtailment of production, as well as the withholding of cotton from the market by the loan operations of the government, brought about a higher price of cotton, and the benefit payments by the government to coöperating planters were large. This is shown in the following table:

FARM INCOME OF COTTON PLANTERS

Calendar Year	*From cotton lint and cotton seed**	*Rental and Benefit Payments***	*Total*
1932	$460,694,000	.	$ 460,694,000
1933	577,551,000	$111,419,000	688,970,000
1934	862,854,000	142,294,000	1,005,148,000
1935	711,786,000	145,910,000	857,696 000

* *The Agricultural Situation*, U S Department of Agriculture, April 1, 1938 p 11

** Nourse, Davis, and Black, *Three Years of the A A A*, 1937 p 586

The price of cotton, which hovered around 11 cents a pound during 1935, did not however reach the parity price of 15 to 16 cents a pound, and there was little indication that the government intended to press very vigorously to bring this about.

Nevertheless despite the success of the AAA in achieving its immediate goal, the cotton program has been subjected to more persistent criticism than has the wheat program or indeed any other program of control. The textile industry was bitter in its criticism and blamed the processing tax, levied at the rate of 4.2 cents a pound on the first processing of cotton, for a decline in the consumption of cotton goods and a consequent accumulation of stocks on which the tax could not be recovered from consumers.

Even more alarm was felt for American exports of cotton, which in the past have amounted to roughly half the crop. While America was reducing its production, other countries were increasing their

own production. Brazil, which in 1932-33 produced a crop of only 448,000 bales, produced 1,712,000 bales in 1936–37. India, the second largest producer in the world, increased its output in these years by about 35 per cent. Almost every country in the world which contributes cotton to world trade showed an increase in production. For the world as a whole the cotton crop increased from a total of something over 25 million bales in 1932–33 to nearly 34 million bales in 1936-37, while the share of the United States in total world production decreased from 55 per cent to 40.6 per cent. Foreign mills were rapidly changing their machinery so as to be able to use cotton from India or Brazil.

Here was a permanent danger to American export trade in cotton. In the five-year period ending on July 31, 1929, the United States exported on the average 8,513,000 bales of cotton yearly; in the year ending July 31, 1934, the exports were 7,534,000 bales; in the year ending July 31, 1935, only 4,797,000 bales were exported, and to date cotton planters have not recovered their former share of the world market, either absolute or relative. At the same time American exports of cotton gins and presses increased materially, in 1934 their value was more than twenty times as large as in 1932, five times as large as in 1933, and twice the average value of such exports in the pre-depression period.

Although the decline in cotton exports is due in part to the growth of a spirit of nationalism in foreign countries, which sought to build up domestic sources of supply, it is in large measure to be attributed to the restriction program and the "loan policy" of the government, the latter functioning, not as a loan, but as a purchase by the government which kept many millions of bales from the world market. The result was the maintenance of world prices high enough to encourage the continuance of plans for expansion of cotton planting in other parts of the world and in addition the maintenance of a differential between the prices of American and foreign cotton sufficient to induce some foreign manufacturers to alter their machinery to permit the use of cotton from sources other than the United States.

Other programs of control. The agricultural adjustment administration embraced other programs of control. Some of these have

already been mentioned briefly, such as the control applied to the production of corn and hogs, sugar, and tobacco, and the marketing agreements, most notable in the case of fluid milk, made with the handlers of a number of products. The government engaged also in a number of other operations which, unaffected by the decision of the Supreme Court declaring the main part of the agricultural program unconstitutional, have continued to the present time and are apparently a regular part of the permanent national program.

The Federal Surplus Commodities Corporation, an AAA subsidiary, was provided with funds, equal to 30 per cent of the customs receipts, for surplus removal operations. In part this consists of the subsidized export of surplus commodities, but in large measure it has involved the purchase of surplus farm products for distribution to people on relief. To some extent funds have been expended to find new uses and new domestic markets for farm products. In the fiscal year ending June 30, 1938, the corporation spent about $54,000,000 to divert price depressing crop surpluses from regular channels and gave over a billion pounds of foodstuffs to state and local welfare agencies for distribution to families on relief. For the year 1938-39 about $80,000,000 is available for this purpose.

The program of the AAA for the purchase of sub-marginal land and its retirement from cultivation is another part of the permanent program. This plan was started in the first years of the AAA and has continued as Congress has made funds available for further purchases. The program at present contemplates the eventual acquisition of something over nine million acres of land unsuitable for agriculture and the diverting of this land to other uses, such as forestry, grazing, wild life, recreation, watershed protection, and flood control. The former owners are to be transferred to different localities and possibly to different occupations.

Direct effects of the AAA program on agriculture. Having observed some of the direct effects of the programs for controlled production as they have been applied to wheat and cotton, it remains to consider the effectiveness of this program of control on the farmers as a whole and on business recovery. The following

table shows graphically the cash income of farmers in each year from 1924 through 1937.[1]

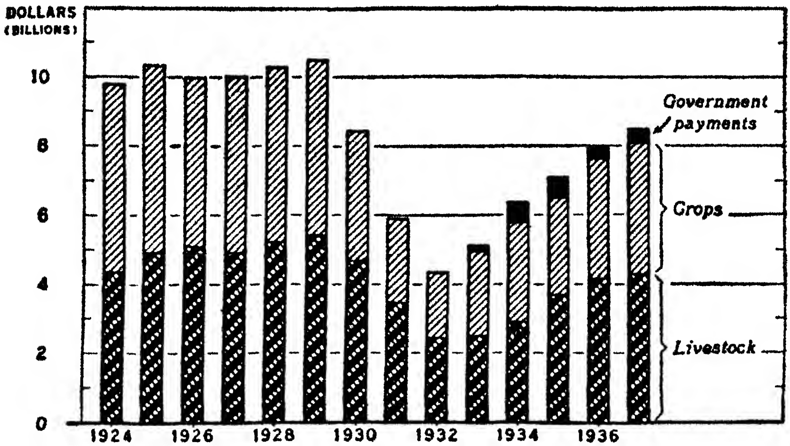

Fig. 7. Income from Farm Marketings and Government Payments to Farmers, United States, 1924–1937[2]

The cash income in 1932 is estimated to have amounted to $4,328,000,000, to $7,090,000,000 in 1935, and according to the preliminary estimate, to nearly $8,500,000,000 in 1937. This increase in income is the result both of benefit payments and of enhanced prices obtaining in the open market. The total amount of benefit and rental payments from the passage of legislation establishing the AAA (May 12, 1933) to December 31, 1935, was 1,108 million dollars. Included in this are rental and benefit payments amounting to 333 million dollars for cotton, 255 millions for wheat, 53 millions for tobacco, 57 millions for sugar beets, 397 millions for corn and hogs, 9 millions for rice, and 2 millions for peanuts. In addition around 225 million dollars was disbursed in surplus removal and drought relief operations, the largest item being 120 million dollars for the removal of cattle from the drought area. Total expenditures for all purposes are reported as amounting to $1,480,057,513.

Agricultural prices rose during this period of control by the AAA.

[1] Though the particular legislation we are now considering was declared unconstitutional early in 1936, its influence naturally extended beyond that date.

[2] Department of Agriculture, *The Agricultural Situation*, Jan 1, 1938, p. 9

From the middle of 1933 to the fall of 1935 farm prices rose on the average by more than 46 per cent, while the wholesale prices of non-agricultural prices rose by 26 per cent. The Department of Agriculture index of the purchasing power of farm products, taken as 100 for the period 1909-1914, which had gone as low as 49 early in 1933, was up to 70 in June, 1933, and averaged 64 for 1933, 73 for 1934, and 86 for 1935. Prices of individual commodities had not, of course, moved at the same rate. Wheat in January, 1936, was 83 per cent of parity; corn 66 per cent; hogs, 98 per cent; and cotton, 71 per cent.

The AAA record shows substantial accomplishment in decreasing the carry-over of certain crops. The wheat carry-over of 393,000,000 bushels in 1933 had been decreased to a normal carry-over estimated at 152,000,000 bushels on June 30, 1935. The world carry-over of American cotton was down to 9,000,000 bales, of which about 5,000,000 bales were in the hands of the government. In the case of corn the drought of 1934 handled the surplus problem with conspicuous success.

For this improvement in the condition of agriculture the AAA cannot of course claim full credit. The farm value of the controlled crops was 85 per cent greater in 1934 than it had been in 1932, but the farm value of the non-controlled crops was 50 per cent larger. A part of this remaining differential in favor of controlled crops is to be accounted for by the peculiarly severe effects of the drought. In the matter of area reductions also the AAA can take only partial credit for the changes effected, for while it is true that the acreage of controlled crops harvested declined 23 per cent between 1932 and 1933, individual adjustments to market conditions on private initiative were sufficient to produce a 13 per cent reduction in these years in the harvested area of non-controlled crops.

The AAA and recovery. As the title of the agricultural adjustment act indicates, the AAA program was designed, not only to relieve the acute agricultural distress, but also as an integral part of the more comprehensive recovery program to benefit the whole nation. In fact the chief justification for the AAA program lay in the argument that, if the farmers' buying power were restored, orders would begin to flow to industrial centers, and employment

would pick up. This in turn would mean a greater demand for farm products and so on until more prosperous conditions generally had been brought about.

In the preceding section we have seen evidence of an increase in farm income, arising partly from improved prices and partly from benefit payments. Unquestionably restored confidence on the part of farmers and improved credit facilities tended to increase their disposition to purchase both consumable goods and producers' goods. There is also plenty of evidence of increased buying in farm areas, of the payment of delinquent loans to local banks and other institutions, of the payment of over-due taxes, and of other aspects of improved financial condition. We know likewise that a measure of industrial recovery is to be found in this period.

The influence of increased farm buying on general recovery is difficult to appraise. The general conclusion of the Brookings report to which we have referred above is that the AAA program had an appreciable and positive effect on recovery. This resulted from an enlargement of total purchasing power (chiefly from government subsidies and loans) and not merely a transfer from one group to another, an acceleration in the rate of turnover of the funds involved, and a considerable advance in the time at which this purchasing power was made available. The Agricultural Adjustment Administration estimated that 40 per cent of the increase in factory employment which had taken place by the fall of 1935 was attributable to improved rural trade. The authors of the report state that "there seems no reason to challenge the approximate validity of the estimate of the AAA," but they offer no quantitative estimate of their own.[1]

Unconstitutionality of the AAA program. On January 6, 1936, the Supreme Court, by a vote of six to three, ruled that the use made of the processing taxes constituted regulation of agricultural production and was unconstitutional in that it invaded rights reserved by the Constitution to the states. The tax was disapproved of as an attempt by Congress to attain ends indirectly which it did not have the power to seek directly.

[1] Nourse, Davis, and Black, *Three Years of the AAA*, 1937, p. 448. See Chapter XIV for a discussion of this general problem.

The immediate effect of this decision was to put an end to the collection of processing taxes, the negotiation of contracts for reducing the production of basic farm products, and the payment of benefits for compliance therewith. Congress did however provide funds during the course of 1936 for the payment of benefits to farmers who had lived up to the terms of agreements entered into with the government prior to January 6. A logical sequel to the court decision was the repeal, at the request of the President, of the cotton, tobacco, and potato control laws.

January 6, 1936, does not however mark the termination of the activities of the Agricultural Adjustment Administration. Not only was new legislation, described below, enacted shortly thereafter, permitting the Administration to seek the same objectives in a different and less effective manner, but many of the powers conferred by the act of 1933 and its amendments were unaffected by the court decision. The Administration still retained the right to negotiate voluntary marketing agreements, to subsidize the export of farm crops, to withdraw marginal land from cultivation, to purchase and dispose of surplus products, to make commodity loans, and to allot sugar quotas.

The soil conservation and domestic allotment act of 1936. This act, approved on February 29, 1936, was passed as an amendment both to an earlier conservation act and to the agricultural adjustment act. Together with the remnants of the adjustment act listed above it provided the basis for the operations of the adjustment administration until the passage of the act of 1938.[1]

The purposes of this act are stated to be (1) the preservation and improvement of soil fertility; (2) the promotion of the economic use and conservation of land, (3) the diminution of exploitation and wasteful and unscientific use of national land resources; (4) the protection of rivers and harbors against the results of soil erosion; and (5) the reëstablishment, at as rapid a rate as the Secretary of Agriculture may determine to be practicable and in the general public interest, of the ratio between the purchasing

[1] An act of 1937 provided loans of 9 cents a pound and a subsidy of 3 cents a pound on a portion of the cotton crop. These payments, according to the act, were to be contingent upon the enactment of new farm legislation, and the coöperation of farmers in a program under this legislation in 1938.

power of the net income per person on farms and that of the income per person not on farms that prevailed during the five-year period, August, 1909, to July, 1914.

The act contemplated coöperation between the federal and state governments in carrying out this program. States were expected to enact legislation permitting the setting up of soil conservation districts within the states, with power to prescribe local land use regulations for the prevention and control of erosion. Up to September, 1937, 24 states had passed such laws. The determination of specific objectives, technical guidance, and payments for compliance was to be provided by the federal government.

In practice the Secretary of Agriculture has made payments to farmers who have agreed to divert land from soil-depleting crops to soil-conserving crops, such as grasses and leguminous crops, and to follow soil-building practices. The program for 1938, as announced by the Secretary of Agriculture prior to the passage of the agricultural adjustment act of 1938, provided for a national goal in terms of acreage for soil-depleting crops and a national objective for soil-building crops and practices. The national goal was to be subdivided into state, county, and individual farm goals. There were individual soil-depleting crop goals for cotton, corn, tobacco, peanuts, rice, and potatoes. Other crops included in the total soil-depleting crop goal were classified together in a general soil-depleting crop goal. A maximum payment was to be calculated for each farm, with payment to be made for keeping within soil-depleting crop goals and for attaining soil-building goals. The 1938 goal for all soil-depleting crops was between 275 and 290 million acres, as contrasted with an average of 305 million acres in cultivation in the period 1928–1937.[1] Save for changes permitted or required by the act of 1938 this program has been carried out in 1938.

Payments to farmers coöperating in this soil conservation program come directly from the Treasury and are raised either by taxation or by borrowing. It is therefore a burden which the taxpayers as a whole have to bear, and that it is not an insignificant burden is evidenced by the fact that in 1936 farmers received 287

[1] *Report of the Secretary of Agriculture, 1937.* Washington, 1937, pp. 12–15.

million dollars as benefit and conservation payments; in 1937 the sum was 367 millions.

The agricultural adjustment act of 1938: Principal features. The soil conservation act has never been considered by the government or the farmers as a satisfactory substitute for the original farm legislation, since it gave relatively little opportunity to control the production of various crops. Its chief feature, aside from the promotion of soil conservation, in which the administration took a genuine and warranted interest, consisted in the improvement of farm income through the payment of subsidies from the public treasury. In the circumstances it was inevitable that an attempt should be made to pass further legislation, which would not be deemed unconstitutional, giving the federal government comprehensive powers of control over agricultural production. This found expression in an act, approved on February 16, 1938, the most important provisions of which are summarized below.

(1) *Soil conservation.* The act continues and supplements the soil conservation program of 1936. The Secretary of Agriculture is authorized to prescribe bases for acreage allotments to carry out the program. Furthermore when Congress provides the funds, payments are authorized to producers of corn, wheat, cotton, tobacco, and rice in such amounts as are necessary to restore the 1909–1914 ratio of farmer's income to the income of others.

(2) *Commodity loans.* The Commodity Credit Corporation is authorized to make loans on agricultural commodities. Such loans are mandatory on wheat, cotton, or corn whenever the stock (as estimated on a specific date) exceeds a normal year's domestic consumption and exports. The corporation must also make loans on wheat if on July 15 of any year its farm price is below 52 per cent of the parity price, and on cotton if on August 1 the average price is below 52 per cent of parity, and on corn if the farm price is below 75 per cent of parity. Loans made to coöperators in the soil conservation program shall be at a rate between 52 and 75 per cent of parity, and loans to non-coöperators shall be at a rate equal to 60 per cent of the rate to coöperators.

(3) *Parity payments.* If and when appropriations are made by Congress, payments may be made to producers of corn, wheat,

cotton, tobacco, and rice to bring the effective return from the commodity up to its parity price.

(4) *Marketing quotas.* The Secretary of Agriculture shall make known his findings with regard to stocks of commodities and shall proclaim a marketing quota if the total supply of any of the five commodities exceeds a given percentage of the supply for a normal year's domestic consumption and export.

No quota shall be effective if more than one-third of those voting in a referendum oppose it.

Acreage allotments shall be made for corn, wheat, cotton, rice, and tobacco on the basis of average production per acre over certain preceding years and shall be so fixed as to produce an amount sufficient to supply domestic needs.

Commodities marketed in excess of individual farm quotas shall be subject to definite penalties set forth in the act for each of the five commodities.

(5) *Crop insurance.* A government-owned Federal Crop Insurance Corporation is established within the Department of Agriculture, with capital stock to amount eventually to $100,000,000. This corporation is authorized, beginning with the 1939 crop, to insure wheat producers against loss in the yield of wheat because of unavoidable causes, such as drought, flood, insects, plant diseases, and the like. Contracts must cover not less than 50 per cent nor more than 75 per cent of the recorded yield for a representative base period. Premiums are payable in wheat or the cash equivalent.

Operations under the 1938 act. In the spring of 1938 Congress provided a sum of $500,000,000 for payments under the soil conservation program and by a later act made an appropriation of $212,000,000 for parity payments to farmers coöperating with the program. Funds also have been provided whereby the Commodity Credit Corporation can make loans on agricultural products, and on July 14, 1938, it was announced that the corporation had set aside $100,000,000 — with more available if necessary — for loans to wheat producers.

Wheat was added to the list of soil-depleting crops, and a goal of 62,500,000 acres was set for 1938, with benefit payments of 12 cents a bushel to the coöperators. The acreage goal was announced

after most of the year's planting had been completed; it had therefore little effect on the actual number of acres sown to wheat, estimated to amount to about 80,000,000 acres. Wheat producers may however qualify for benefit payments by diverting other land from soil-depleting crops to soil-conserving crops and by following soil-building practices. The loan program for wheat, announced in July, 1938, was mandatory according to the law, because the official estimate of the wheat crop, made on July 1, indicated a crop in excess of a year's normal consumption and export. Loans were offered at farm rates averaging between 59 and 60 cents a bushel, the minimum rate permitted by the act, with maturities of seven months for wheat stored in commercial warehouses, and maturing on May 31, 1939, for wheat stored on the farm. For the 1939 crop an acreage allotment of 55,000,000 acres, the minimum permitted by law, was set. Announcement was made of a material increase in the benefits to be paid to farmers who cooperate with the 1939 program. In comparison with a payment of 12 cents a bushel under the 1938 program, something between 26 and 30 cents a bushel was to be paid in 1939. If the benefit payment were set at 30 cents, a cooperating farmer whose allotment was 60 acres would receive a total benefit of $216, provided his average yield in the past had been 12 bushels an acre. Finally the AAA decided to buy wheat up to 100,000,000 bushels at American prices to sell in foreign markets for what it would fetch.

Early in 1938 the Secretary of Agriculture's proposal of a cotton marketing quota, defined as the product of 26,300,000 acres, was approved by the vote of an overwhelming majority of cotton planters. Official reports estimate that 26,904,000 acres, the smallest area since 1900, were actually planted to cotton. Planters coöperating in the soil conservation program and living up to their individual acreage allotments were to receive a benefit of 2 cents a pound. Those who exceeded their allotments — cooperators and non-coöperators alike — were to be fined 2.4 cents a pound on the excess marketed. A loan program was announced also for cotton. Acreage allotments were set for corn, tobacco, and rice, and in the case of tobacco marketing quotas were set after a referendum of tobacco planters.

Comparison of the acts of 1933 and 1938. Perhaps the most conspicuous point of difference between these two acts is that the 1938 act omits the processing taxes and in their place provides that benefit payments are to be paid when and if Congress provides the funds. The question as to the extent to which Congress wishes to subsidize agriculture is thus presented annually. Congress can if it chooses put an end to the entire program by the simple expedient of failing to vote the necessary funds.

In the second place, income parity is set up as the ultimate objective in place of price parity. Price parity is still to be used as a measuring rod, but the goal now is to secure, through a rise in the prices of agricultural commodities and the payment of benefits, a net income for the farmer which will bear the same ratio to the income of the non-farmer that it had in 1909–1914.

A further objective of the 1938 act is best expressed by the term "ever normal granary," which has been so stressed by the Secretary of Agriculture in recent years. This concept implies a steady flow of farm products to the market at prices which are reasonably stable over the years and sufficiently remunerative to the farmer to achieve the major goal of the program; namely, income parity. The acreage allotments are supposed to be set at a figure which will produce the desired quantities of the products, but these allotments are binding only on those farmers who choose to coöperate with the government. Therefore marketing allotments are provided, which in the event of embarrassing surpluses make the marketing of produce in excess of the allotments subject to fines. Marketing allotments are compulsory acreage allotments, affecting all farmers when ratified by a two-thirds vote of the farmers concerned.

The administration recognizes however that weather conditions, insect pests, human frailty in living up to quotas, and indeed errors in estimating accurately both the requirements for consumption and the acreage necessary to produce that amount may result in crops which are overabundant in one year and deficient in another. The commodity loan is specifically intended to even out discrepancies between the requirements for consumption or export and the quantity produced in any given year. In time of glut and low

prices the government is prepared, through the Commodity Credit Corporation, to make loans to farmers on produce stored on the farms or in commercial warehouses. The withdrawal of these stocks from the effective supply will therefore raise prices -- presumably to or near the desired level. When a poor crop comes along and prices would otherwise be higher than desired, the produce in storage is to be released, thus adding to the effective supply and giving assurance of normal quantities for consumption and normal prices.

Crop insurance is another innovation which is expected to work in the same direction, although it is intended also to relieve the farmer of some of the risk which has traditionally been his. Premiums are payable in wheat, the only crop for which insurance is at present provided, thus to some extent decreasing the quantity on the open market. If there is no occasion to pay out benefits the wheat remains in storage, but if a drought or a plague of insects destroys a portion or all of a crop the stored-up wheat will be released for the payment of insurance benefits, giving the farmer a substantial portion of his normal income and assuring to the consumer supplies of wheat at prices close to normal. Thus we have the twentieth century application of the moral of Joseph and the famine!

Nevertheless despite these and other differences in detail, the two acts are essentially similar in spirit, in objective, and in the far-reaching powers of control conferred on the administration. The law indicates the intention of subjecting agricultural production to a degree of regimentation heretofore not considered either possible or desirable

Appraisal of the agricultural program: General principles. The general propositions which, in the minds of its sponsors, justify legislation of the character which we have been studying may be summarized very briefly, as follows:

(1) American farmers can produce more of our basic agricultural products than can be sold in the domestic or foreign markets at reasonably remunerative prices.

(2) Under conditions of free competition surpluses over and above the requirements for domestic consumption and foreign

markets are certain to appear — not in every year, but frequently enough to depress prices to such an extent that farmers fail to secure their fair share of the national income.

(3) Government is able, by appropriate measures of control, to maintain equilibrium.

(4) It is the duty of government to devise and carry out a program which will maintain equilibrium in agriculture.

A permanent program of governmental control of agriculture finds its justification therefore in the belief that competition is unable, even in the long run, to achieve a balance between supply and demand. Has competitive price in fact been a failure as a regulator of agricultural production?

Of the failure of competition to produce equilibrium in the short run history will give plenty of instances. In pre-war years periods of agrarian distress and of agrarian agitation for assistance were not uncommon, but that a reasonably satisfactory balance was secured in the long run is implicit in the fact that the period 1909–1914 is taken as the norm both as to farm prices and farm income for the future.

In the decade following the World War the lack of equilibrium in agriculture was evident enough, but one should bear in mind that at the close of the war agricultural producers were faced with the necessity of making an adjustment of formidable dimensions and that agricultural adjustment is a slow process. This period was marked furthermore with recurring crises, both political and economic, in various parts of the world, which still further complicated the situation. In the United States it was a decade of political agitation for relief for the farmers. In these years Congress passed two important agricultural acts which were vetoed by the President and finally in 1929 enacted the agricultural marketing act.

The times called for the elimination of the submarginal producer and the abandonment of submarginal land, but with the prospect of government assistance throughout this period it is hardly surprising that the submarginal producers hung on to their land as long as they could. Since 1929 there has been little inducement for the submarginal producers of at least the major crops to abandon

cultivation; the 1929 act sought to raise prices without attempting to limit production, while the subsequent acts have paid benefits to intramarginal, marginal, and submarginal farmers without distinction.

There are many critics of the farm program of the government who still have faith in the regulative power of a free competitive market and who believe that the sponsors of government control have condemned competitive price too hastily and on false theory and untrustworthy evidence. These critics would not deny that there was need for some government assistance in 1933, but they believe that governmental action should have been directed primarily toward assisting competitive price to bring about equilibrium more quickly.

There is in fact little ground for belief that the AAA program works in the direction of restoring to agriculture a stability which will enable it hereafter to go ahead without government aid. When benefit payments are made to high-cost and low-cost producers alike the result will be that either poor land or poor farmers or both are kept at work. It is true that a beginning has been made in the purchase of lands which are submarginal for agricultural purposes and their diversion to other uses, but this can have little effect in the present or the near future. Despite the improvement in farm prices and farm incomes there is no assurance that an equilibrium in agriculture has been attained or will be attained which would prevent a recurrence of conditions obtaining in 1933 if government control were abandoned. The very economic weakness inherent in the program of the AAA thus serves as the strongest political argument for its continuation.

The theory of parity income. The ultimate objective of the farm program is to increase farm income up to that point where the net income per person on the farm bears the same ratio to the net income per person not on the farm as it did in the period 1909–1914. This concept of parity income obviously appeals to its prospective beneficiaries; it also appeals to many others as a fair division of the national income — as according simple justice to the farmers.

At the outset one may ask why the division of income existing

in these years should be perpetuated. In the one hundred and more years of national life agriculture has shared in the national income in varying proportions. What makes the period just before the World War so peculiarly representative of justice to the farmer? If it be today representative of justice according to some preconceived standard, will it be tomorrow?

It may be argued that as a practical matter the government must have some standard if it is to attempt to raise the prices of farm products and the incomes of farmers. This is undoubtedly true, but the dangers inherent in the notion that justice in prices or incomes ought to be or can be achieved by government action are not thereby dispelled. If there is a just price for wheat or a just income for a wheat producer, is there not quite as much reason to suppose that there is a just price or a fair income for the producers of every other product or service, and do they not have as much right to ask for justice as the wheat farmer? And if they can muster as much political power as the wheat farmers, will they not get the legislation they ask? There were seven basic commodities enumerated in the original farm act of 1933, but amendments had more than doubled the list before the act was declared unconstitutional. The agricultural adjustment act of 1938 provides parity payments for five major products only, although conservation payments are available to all farmers, but it is hardly reasonable to assume that other products will not be added to the list.

Once the notion becomes accepted that it is the duty of the government to safeguard producers from loss or from the consequences of their own judgment and management, constant pressure by interested groups to secure benefit payments will lead inevitably to an extension of government control. There is no limit to this development short of the socialistic or totalitarian state. There may indeed be those so convinced of the futility of competitive price as the regulator of production as to welcome this. But he who believes that there still is virtue in competitive price, even though it may work hardship at times on the individual, will view the just prices and the just distribution of national income imposed by government with at least as much suspicion as he does the monopoly prices and monopoly profits of private producers.

Subsidies and loans. It appears impossible to avoid the conclusion that the payment of subsidies from the public treasury to impel farmers to reduce the area devoted to the raising of crops has these effects: it necessitates the levying of taxes on consumers as a whole for the benefit of a special class, it serves as a precedent for similar treatment of other groups than the farmers, and it affects adversely the energy, initiative, and self-reliance of the farmer himself.

Experience with government loans on crops in the past does not seem to warrant such optimism regarding this phase of the present program as the advocates of the "ever normal granary" exhibit. On several occasions in the past five years loans have been made on cotton, with the result that the government is at present holding from the market some 10,000,000 bales, with no indication as to when it can be disposed of. Government loans, especially if made without recourse, as has been done in the case of cotton loans, are apt to be the same thing as government purchases. If a time comes when the price is considered high enough to justify the disposal of the commodity, will it be politically expedient to do so? There is sure to be an outcry raised against spoiling what would otherwise be an unusually favorable market.

Conclusions. The federal administration was presented in 1933 with an acute agricultural emergency. From both a political and a humanitarian point of view it seemed necessary to do something for the farmer and to do it at once. One may not believe that even as emergency legislation the program in 1933 represented the wisest choice, but he may recognize the sincerity of the effort to find an immediate remedy for the situation, and he must recognize that the economic position of the farmer has greatly improved since the spring of 1933. However while there is little profit in criticizing legislation which was frankly of an emergency character, the prospect of the development of this program into a permanent policy is another matter. The payment of agrarian subsidies, the loss of export markets through artificially high prices, the possibilities inherent in government crop loans — these and other phases of the program are serious matters. The theory of parity income is not only of doubtful validity itself, but its practical application is

attended with so much danger of an extension of the scope of government control as to constitute a real menace to our present form of economic and political life. Government control of production evokes the spectre of a vast bureaucracy engaged in planning production in ways not essentially dissimilar from those followed in soviet Russia.

EXERCISES

1. Appraise each of the following proposed "solutions" to the farm problem:
   (*a*) Acreage restriction obtained through benefit payments to farmers for withholding land from cultivation.
   (*b*) Government purchase and withdrawal from use of sub-marginal land.
   (*c*) Higher American tariff duties on farm products.
   (*d*) Lower American tariff duties on manufactured goods.
   (*e*) A system of export subsidies for American farm products
   (*f*) Coöperative marketing by farm producers.
   (*g*) Farm mortgage moratoria, or a general scaling down of farm mortgage indebtedness.

2. Why should not the principle of parity prices be extended to other industries? Would it be desirable to seek to restore the pre-war relationship between the prices of automobiles, or of gasoline, and the prices of other goods?

XLVI

REGULATION OF COMPETITIVE INDUSTRY

Present significance of the NIRA. The program of the United States government in 1933-1936 with respect to the control of agricultural production and agricultural prices had its origin, as we have seen in the chapter immediately preceding, in a belief that the operation of the forces of demand and supply cannot be relied upon in a period of serious maladjustment to restore equilibrium in production or to fix prices which are just to the farmers. The continuance of this program into the future would imply doubt of the ability of price adequately to regulate agricultural production at any time. This lack of confidence in the operation of the competitive system is seen also in the efforts of the Roosevelt administration to extend the scope of government control to cover industry as well, with the avowed purpose, not of checking the growth of monopolies or of regulating such industries as are inevitably monopolistic, but of sparing society the abuses of competition and of saving industry from self-destruction.

The legislative measures for putting into effect this program of industrial control were contained in the national recovery act of June, 1933 (NIRA). This act was short-lived, being declared unconstitutional by a unanimous decision of the Supreme Court in May, 1935. While the decision put an end to the NRA as an agency of control, the study of this measure and its brief operation has far more than historical interest. The doctrine that the system of private capitalism functions at times very imperfectly and develops abuses requiring correction at the hands of government is not a novel one nor one which is rejected by even the most conservative, but that competitive industry needs or that society can profit from the amount or type of control authorized by the national industrial recovery act is distinctly new to American thought. The experience under the NRA presents an unusual opportunity to investigate important questions of principle which are far broader in scope and

more enduring in time than this particular undertaking. Moreover the influence of code control on industry cannot be thrown off at once, and we still have the NRA with us, in miniature at least, in the control at present exercised over the oil and bituminous coal industries.

The present chapter is therefore devoted chiefly to the study of the NRA, in order to examine its economic philosophy and appraise its practical achievements both in promoting recovery and in reforming business practice.

Objectives of the NIRA. The national industrial recovery act was distinctly emergency legislation, being enacted for a period of two years from June 16, 1933, and having as its foremost ostensible objective the promotion of industrial recovery. A second objective was to secure reforms in business practice in order to eliminate the ill effects of excessive or predatory competition and to achieve an equilibrium in business activity that might make industry in the future less susceptible to disturbance than it had been in the past. The act declared that, because the flow of interstate commerce had been diminished by the existence of a national emergency, it was the duty of Congress to encourage the flow of commerce and to promote the national welfare by providing for the organization of industry into trade groups for the purpose of coöperative action. Internal industrial coöperation, or what has been frequently termed "self-government of industry," was intended (1) to secure united action of labor and management under government sanction, (2) to eliminate "unfair competition," (3) to promote the fullest utilization of the productive power of industry, (4) to avoid undue restriction of production except as temporarily required, (5) to increase the purchasing power of consumers, (6) to reduce unemployment, (7) to improve standards of labor, (8) to rehabilitate industry, and (9) to conserve natural resources. In practice recovery was sought primarily by spreading employment through shortened hours of work and by increasing industrial payrolls through higher rates of pay. These labor provisions for recovery were supplemented by provisions for the maintenance of prices at levels sufficiently high to permit firms to meet the increased labor costs. The second phase of the program became of increasing importance as the failure of the

labor provisions to promote recovery became inescapably apparent.

Principal features of the NIRA. The national industrial recovery act gave the President and his administrators wide powers to promote self-regulation of industry under federal supervision as a means of getting the joint program for industrial recovery and reform into operation. It was provided that voluntary codes of fair competition, drawn up within each industry by the industrial groups, should be submitted to the federal government for correction and final approval. The government was given power to enforce the provisions of the codes adopted and to assess fines for their violation, to draw up and impose a code for any industry not submitting a satisfactory one of its own, and for the space of one year to use a system of licenses for operation in business so as to force any unwilling minority of an industry into line. The licensing power was allowed to lapse, leaving persuasion and the threat of legal prosecution, with penalties of fine and imprisonment, as the weapons for forcing compliance. Codes and agreements under the act were endowed with exemption from the operation of the anti-trust laws, but monopolistic practices remained specifically under prohibition.

Labor was granted the right to bargain collectively through representatives of its own choosing, and the "yellow dog" contract — *i.e.*, a contract which binds the worker, as a condition of employment, not to join a union -- was positively outlawed. The President might prescribe maximum hours of work and minimum rates of pay if he did not care to approve those suggested in a code. The President might use licenses, quotas, or an embargo to limit imports that threatened the wage and price objectives of any code.

The oil industry was singled out for special mention, with the grant of three specific powers to the President: he might start proceedings with the Interstate Commerce Commission for the control of pipe lines and the fixing of reasonable rates; he might start proceedings to divorce from a holding company any pipe-line company tending to create monopoly by its unfair practices or exorbitant rates; he might prohibit transportation across state boundaries of any petroleum or its products produced or withdrawn from storage

contrary to state regulations (known as "contraband oil" or more popularly as "hot oil"). This oil section of the act met judicial disapproval (on January 7, 1935) as an improper delegation of discretionary power and was replaced by other legislation.

Code making. The President set up the National Recovery Administration to get hearings started on proposed codes and to proceed with the general administration of the act. Most proposed codes were drawn by trade associations or similar organizations, and code hearings were intended to ascertain whether the proposals were sufficiently representative of the opinion of the industry, whether labor provisions were properly inserted, whether provisions with respect to fair practices were properly non-discriminatory, and whether the interests of the consumer had been safeguarded. Industrial, labor, and consumer advisory boards were formed to give assistance at this stage of the proceedings, but in drafting any code the dominant members of an industry were naturally in a preferred position for protecting their particular interests in the code finally approved. The act became a law in June, 1933; by December some 170 codes, covering most of the major industries, had been approved, and about 700 others were in the process of drafting or hearing The President estimated at this time that 70 per cent of all workers in industries eligible for "coding" were working under permanent codes; the bulk of the remaining workers, in the smaller industries, were employed under the so-called "blanket code."

In order that some of the effects on employment and wages which the NIRA intended might be realized during the period while the codes were being drawn up and approved, the President in July, 1933, invited industry to sign at once the "President's Reemployment Agreement," more popularly known as the "blanket code." The signers of this voluntary agreement were bound to put into effect at once certain provisions regarding maximum hours of work, minimum wages, and child labor until approved codes for their respective industries were formulated, or until January 1, 1934. At this date however there had been so many delays in forming codes that it was found expedient to extend the life of the "blanket code" until May 1, 1934.

The taking of signatures of individual employers on this agreement was handled through the Post Office Department. The wholesale signing of the agreement represented voluntary coöperation by individual employers, stimulated by the psychology of patriotic appeal and under a more or less concealed threat of public boycott for non-participants. This was the flight of the Blue Eagle, the symbol of a gentleman's agreement with the President, later superseded in coded industries by the Code Eagle, a symbol of compliance with the law.

The NRA and business recovery: "Spreading the work." The complete recovery program of the administration was a highly diversified one, embracing among other things public spending on a huge scale, inflation, and an increase in farm income. The labor provisions of the codes were expected by the sponsors of the NRA to be a major factor in this program. Business recovery was to be achieved, partly from the diffusion of purchasing power among the greater body of employees which it was hoped a shorter working day would entail, but in the main from the increased consumer purchasing power which would flow from the employment of a greater number of workers at higher rates of wages.

The first attack on the problem of recovery is a variant of the "spread-the-work" doctrine, which attracted much favorable attention in the early days of the depression, when many employers gave concrete expression to their belief in it by placing their working forces on part-time employment at the same hourly rates of pay. In this guise it is a method of placing the burden of unemployment more evenly on the shoulders of the working class. Workers who otherwise would enjoy full-time employment have to share their jobs with others. For the accomplishment of this purpose in an emergency, this device has merit. It does not however offer any hope for promoting recovery, save on the dubious assumption that the larger group of employees on part-time would spend a greater portion of the same total wages than would a smaller group on full-time. Indeed there are grounds for believing that the employment of two shifts might decrease the effectiveness of the producing unit and thus bring about a rise in costs, a decreased volume of sales, and a decrease in the total demand for labor.

Wages and hours. The main hope for recovery from the NRA program lay however in the provisions for minimum wages and maximum hours of work, which were expected to bring about the employment of a larger number of men at higher rates of pay than formerly and consequently to cause an increase in the total money wages. It was expected that production would increase and that there would be an increased money income with which to purchase these added goods. At first this new demand would be focused on the industries making consumers' goods, but eventually it would be reflected in industries devoted to the production of capital goods. Once this program was under way, it was anticipated that the process of expansion would proceed automatically.

The business man was asked, as a patriotic duty and in enlightened self-interest, to assume greater labor costs at a time when profits were negative, non-existent, or precariously small, with the assurance that sometime in the future his outlay would be more than repaid to him through an increase in the demand for his product. He had however a number of alternative paths. It was within his power to avoid a larger wage bill, either by reducing his laboring force, or by reclassifying his laborers and decreasing the wages of the more highly paid labor sufficiently to compensate for the increased cost elsewhere. If he chose the first of these alternatives, we would find decreased production, increased costs of production, and higher prices, but the same money income in the hands of the workers. Clearly this would afford no stimulus to business activity; indeed we should expect it to have a depressing influence. If the business man followed the second plan, production would be decreased somewhat because of the shorter working day or week, and prices would have to rise because of this. Again there would seem to be no stimulus to business.

On the other hand, industry could increase the total wage payments by paying the same force higher wages on the average or by increasing, as was hoped, the size of the labor force. If the former labor force were retained, costs of production would inevitably increase, not only because of the higher wages paid but because of the reduced scale of production which a shorter working day would probably entail. Cost of production would consequently increase

more rapidly than wages; expressed differently, the increased money wages would not be sufficiently large to purchase even the decreased output at any price sufficient to cover all the costs of production.

There was however still another alternative. Suppose that industry were to engage more men, and enough more not merely to compensate for the shorter working day but to produce more goods, and were to pay them higher average wages. In this case one could expect an effective increase in demand for the products of industry, provided the increase in wages were actually spent for goods or services and not used to pay off old indebtedness or for hoarding, and — even more important – provided that prices were raised by an amount less than proportional to the increase in labor costs. In periods of depression most plants have much surplus capacity, which results in a high average cost of production. Were production to be expanded, the burden of fixed costs which each unit bears would be decreased, with the result that prices might be raised by enough to return to industry the full amount of the extra wage payments and yet not be increased in proportion to the increase in labor costs. If this were true, and the increased wages were not simply a deduction from profits which otherwise would have gone to stockholders, but came from profits which might have been used to pay off bank loans or to build up idle bank balances or were to be met by borrowing from banks, there would be not only an increase in the number of dollars available for consumer spending, but an actual effective increase in demand for goods.

Why did NRA fail? There was therefore a chance for the labor program of the NRA to succeed in promoting recovery. Is its failure to do so to be ascribed to incompetency in business leadership? Was the recovery program wrecked by business leaders who had as much to gain from it as any group but were so lacking in constructive imagination as to fail to see how they could gain? Such is the verdict of some. But it should be remembered that it is difficult indeed to make a careful business man who is already operating at a loss or on a very slight margin of profit see any virtue in going into debt to employ more men at higher wages. The reward seems to him too intangible, too vague, and too far into the future, and the

danger of producing unwanted and unsaleable goods too immediate. Actually any one of the first three alternatives described above seemed a more prudent path to follow than the fourth, which alone offered any hope of promoting recovery.

There was moreover one other factor, an inconsistency in the NRA itself, which further complicated the situation. Code control afforded business men an opportunity to control prices as well as objectionable competitive practices, and the temptation to restrict output and raise prices was too much for some to resist. Monopolistic tendencies set in at a fairly early date and sounded the death knell of any hope of recovery in the industries where it was found.

With respect to the possibility of promoting recovery through the labor provisions of the NIRA our conclusion is to this effect: In the first three cases there appears no possibility of any stimulation to business recovery; in fact the opposite is more likely to be true. In the last case there is the possibility, but the actuality depends on the price policy of industry; if it were to make increased production and low prices its goal, the program might have a chance of success, but if industry were to seize the opportunity which the code controls afforded and were to raise prices more rapidly than labor and other costs increased, the program would be doomed to failure. The trend toward monopolistic prices made the failure of this part of the program of the NRA perhaps more dramatic, but even in the absence of monopoly there would seem to be little reason to expect the labor program of the NRA to result in an effective expansion of consuming power in the hands of labor.

Reduction of wages and recovery. The failure of business leaders to respond to the plea of the NRA to employ more men at higher wages was the result not alone of their failure to see any immediate personal advantage in so doing, but also because many doubted the practical validity of the reasoning given above. Their business experience told them that an important immediate cause of their difficulties was that wages were in many instances already too high in relation to other prices. Their conviction was that wages needed to be reduced rather than raised.

A policy of reducing wages has been attacked as a policy of futility, a gesture of despair, a relic of the "horse and buggy" era.

There may be better ways of meeting the problems of depression, but it must be recognized that this is one way of attaining an end that must be reached before depression is over. We are operating on a system which makes the confident anticipation of profits a necessary prerequisite to business expansion. Profits are conditioned upon a balance between costs and prices, and in periods of crisis and depression these get seriously out of alignment. Some costs are fixed by contract, and others move very sluggishly. The rate of wages is one of the more slowly moving costs of production, but if the profit margin is to be restored, wages must come into equilibrium with other costs and with the price of the product.

Attacks on the policy of reducing wages come partly from those who believe it contrary to social justice. They point out that many other costs — interest for example — are the result of contract and therefore to a great extent immune to deflationary influences; it seems to them manifestly unjust that labor should be called upon to bear the brunt of depression. It is also criticized by those who believe that a more profitable avenue of approach to the general problems of depression is to attempt, by what some choose to call "reflationary" measures (currency manipulation, government spending, and the like), to raise prices. This means of course a decline in real wages if monetary wages are kept stationary. The merit of the plan, if successful, lies in the psychological stimulus which comes from rising prices and also from the fact that the burden of debt is made relatively easier to bear.

It is not our purpose to engage in this controversy or to examine the merits of alternative plans of meeting the problem of disequilibrium between cost and price, but merely to assert the plain hard truth that, until the prospect of profit improves, either through a reduction of costs or through an increase in other prices induced by methods which do not restrict output, industry cannot be said to be even on the road to recovery. We recognize the practical and political difficulty which would confront any administration which had the courage to announce such a policy, but it is not without the bounds of possibility that it might meet with less antagonism from labor leaders than is commonly thought, provided there were coupled with it adequate assurance that wages would rise in a semi-

automatic manner as soon as business conditions permitted. It is further not only possible but quite probable that labor as a whole would gain, both in money income by increased employment at the lower rates of pay and in real income because of the general reduction in prices. Society itself has more than a passing interest in such a concession to secure the support of labor, because of the fact that a more rapid adjustment in periods of rising prices between wages and profits than is customary would tend to remove one of the causes of the disequilibrium which leads to crises.

The labor provisions and reform. The emphasis placed on the recovery aspects of the labor provisions of the NIRA should not lead us to neglect the fact that minimum wages and maximum hours of labor are considered by many as desirable objectives quite apart from any influence which they may have on business recovery. It is a well known fact that some plants have been able to keep going during the depression only by the exploitation of labor; *i.e.*, by sweatshop wages and hours or by the employment of child labor at low wages. It is equally true that the lowest paid wage earners in some industries have not secured what many consider the minimum necessary to support life on any decent basis. To social reformers, oppressed by these conditions, the inclusion in the codes of provisions for minimum wages and maximum hours represented a distinct achievement in social welfare. They believe that "parasitic industries," or industries unable to pay the minima laid down, should be forced out of existence and their custom and labor transferred to the more efficiently managed industries.

This is not the place for a discussion of the merits of measures of social reform, but it may not be out of place to point out that many who are wholly sympathetic with plans for improving the life conditions of those on the fringe of existence feel that a period of black depression is an inappropriate time in which to launch far-reaching schemes of social reform. Business recovery and reform may be antagonistic objectives.

The provisions in codes to eliminate child labor from industry may be viewed from the point of view of reform or of recovery. As a measure of reform they represented an achievement which had been sought in vain for many years by social reformers. As a

method of promoting recovery they were at best of negligible importance. The presumption would be that the work previously performed at low wages by young people would henceforth be done by adults at higher wages. Some increase in money income going to this group might therefore be expected, but its influence on recovery would as in all cases depend on the price policy of industry.

Section 7a of the NIRA, providing for collective bargaining, should also be considered primarily as a measure of reform. In so far as the application of the provision did actually result in increased unionization and the strengthening of the bargaining power of labor, it would tend to increase labor resistance to wage reductions and therefore make more difficult the readjustment of wages where they were out of line with other costs and prices.

Monopolistic tendencies of the NRA. The anti-trust status of code practices had not become clear up to the time of the demolition of the whole code system. The national industrial recovery act specified that codes were exempt from the anti-trust laws but also made the absence of monopolistic intent in a code a condition for Presidential approval of the code. The President announced in an executive order that codes could not shield price fixing and discrimination, that the Federal Trade Commission or the Department of Justice would investigate complaints of monopolistic practices, and that the anti-trust laws were still in force.

The interest of industry in the control of prices was evidenced by the frequency with which direct and indirect attempts at price fixing were included in the codes. Four principal approaches to the goal of higher profits through price controls are to be found in the more than 600 codes finally placed in operation: (1) the fixing of minimum prices, (2) prohibitions against selling below costs of production, (3) open-price arrangements, and (4) limitations on production or on productive capacity. The power to fix minimum prices represented in an extreme form the rejection of the competitive method of price making, whether the "minimum" was defined as "the lowest reasonable cost of production," or the "fair and reasonable" cost, or the cost of the "lowest cost representative firm." Prohibitions against sales below individual costs, except to meet legitimate competition, were included in 352 codes, and in

some instances codes were rendered even more restrictive of competition by the requirement of uniform and arbitrary allowances for certain elements of cost. In addition, over 100 codes contained some formal prohibition against destructive price cutting. The NRA permitted the establishment of over 400 "open-price plans," under which all prices and any changes therein were filed with the code authority and published to all members of the industry; 297 of these plans provided for a "waiting period" before any price change should become effective. In origin the open-price arrangements were intended to discourage destructive price cutting, but they served equally effectively as a means whereby the disapproving pressure of the trade or industry could be focused on any nonconforming member. There is little doubt that such price provisions, together with requirements for classifications of customers, uniformity in terms of sale, and reporting of orders, as administered by code authorities frequently dominated by the larger producers in an industry, worked in the direction of a restraint on competition and a raising of prices beyond the point necessitated by increased labor costs.

The allocation of quotas of permitted production, the limitations placed on the total hours that machines or plants could be operated, and the restrictions on the installation of new machinery worked in the direction of decreased production and higher prices. At the same time it tended to keep in operation the high-cost inefficient producers who could not otherwise have survived. In a critical attack on the code method of governing industry one commentator said, "They are permitted under the codes to restrict output, to limit the introduction of machinery, to add to the price of goods freight charges from certain arbitrarily established shipping points, although the actual shipment to the consumer may be from points much nearer. But more important than all this, they have the right to set up machinery to act in concert and collusively without adequate public control. This is called self-government in industry, the corrective to a competition which was proving so unprofitable. Another name for it is monopoly." [1]

[1] From a speech by Professor A. Anton Friedrich of New York University at a conference in Washington on the NRA, *New York Times*, Feb. 28, 1934, p. 1.

Retreat of the NRA. In the summer of 1934 a National Recovery Review Board, set up specifically to inquire into monopolistic tendencies of the NIRA, issued three reports, known as the Darrow reports. The general conclusion of these reports was that the codes of a number of leading industries, if not all, were controlled by the representatives of the larger firms, favoring and producing conditions of monopoly, and seriously injuring small businesses. In June, 1934, the NRA announced a general change in principles which, while not immediately interfering with existing codes, sought to check more vigorously than before the monopolistic tendencies in future codes These developments, particularly with reference to price policy, betokened a considerable retreat from the objective of helping to stimulate business recovery by arbitrarily holding industrial prices in alignment with costs of production, and it betokened also something of a revival of the economic philosophy of the anti-trust laws.

At the same time practical considerations had forced upon the NRA at a rather early date some narrowing of the scope of code application, which had been intended originally to be practically universal. In the spring of 1934 an executive order of the President exempted small enterprises in towns of less than 2,500 population not in industrial areas from compliance with wage, hour, and price provisions of codes and from liability for assessment for support of code authorities. A little later the NRA rescinded the code regulations of fair practices and prices in seven "service" industries, including the cleaning and dyeing industry, where particularly great difficulties had been encountered in trying to enforce these national regulations. National provisions as to wages, hours, collective bargaining, and child labor remained in effect, and the NRA gave recognition and support to any local code of fair practice approved by 85 per cent of the members of a "service" industry in the particular locality. These "service" industries represented purely intrastate business, in which national regulation had no legitimate place. The NRA, without directly admitting any limitation of its own power, had previously urged state governments to give code regulations the force of state law, but nothing was accomplished on this line. In the spring of 1935, the administration was

indirectly shaping its policy toward removal of compulsion generally in the case of all purely intrastate business. This question soon reached an emphatic settlement at the hands of the Supreme Court.

The question of constitutionality. From the earliest days of the NRA there was grave doubt as to the constitutionality of the legislation upon which it was established. This current of disbelief militated against the effective enforcement of the codes pending a final test in the courts. The later drift of government policy in the direction of a narrower application of code control reflected in part an appreciation of the probable unconstitutionality of the system. The chief hope of the "New Deal" administration seems to have been that by the declaration of an emergency the legislation might be considered constitutional. Numerous decisions on cases in the lower federal courts had indicated a preponderant judicial opinion that the NRA constituted an improper application of federal power. Appeals by the government on cases decided against it were notably lacking. From one case appealed to the Supreme Court the government withdrew before argument. There was obvious reluctance on the part of the government to submit the NRA to the final test of legality, but the test had to be undergone at some time if any semblance of code enforcement was to be preserved.

Supreme Court decision. As a test case of the constitutionality of the recovery act and the codes adopted thereunder, the government finally selected the so-called Schechter case, arising from alleged violations of the live poultry code for metropolitan New York. At the end of May, 1935, the Supreme Court delivered its unanimous opinion announcing a clean-cut determination of three important matters. (1) The Court declared that the code-making authority conferred on the President by the recovery act was an unconstitutional delegation of legislative power. This authority was in effect sub-delegated to trade associations to write laws to govern themselves. The opinion was based on the facts that the act did not simply foster voluntary effort for achieving a perhaps necessary increase in business coöperation, did not seek merely to endow voluntary business groups with privileges or immunities, did not supply standards for business activity nor prescribe rules of conduct to be applied to particular conditions as determined by

appropriate administrative procedure, but authorized the making of codes through executive action, with all the force of law and binding equally on those who assented and those who did not, in order to accomplish these legislative ends. (2) The Court held further that the federal government did not have the right to regulate wages and hours of employees in intrastate business merely because of their indirect effect on interstate commerce and that the contention of the government in this respect pushed too far the doctrine of conditions "affecting interstate commerce." (3) It was the opinion of the Court that no economic emergency could justify breaking down the constitutional limitations on federal authority.

Remnants of the NRA. This decision of the Supreme Court merely removed some lingering uncertainty as to the time of the eventual demise of the Code Eagle. The enforcement of code provisions in most industries had been declining in effectiveness for many months, as continuing violations, particularly of trade practices and price regulations, had gone unpunished and apparently unpunishable. In the absence of any clearly valid police power, the force of economic circumstances was rapidly strangling the NRA. The Supreme Court pronounced the obituary. All of the nearly 600 codes in force became immediately inoperative, and all enforcement activities were suspended. Some retail price cutting occurred, some hours of work were increased, some wages were reduced. Some industries reported a general increase of "chiseling" practices. In general there was little immediate effect from the collapse of the code system, as most of the now forceless standards continued to be observed about as fully as they had been observed before. Congress turned its attention to the immediate salvaging of as much as possible of the NRA. To this end a resolution was passed extending a skeletonized NRA for a period of nine and one-half months to April, 1936. This resolution provided simply that business might enter into voluntary agreements for guaranteeing collective bargaining to labor, setting minimum wages and maximum hours of work, abolishing child labor, and prohibiting unfair trade practices of the type already forbidden by statute; it specifically required that no trade practices should be adopted which were illegal under the existing anti-trust laws. The NRA continued for this limited

period as a statistical investigative body, designed partly to give publicity to any "chiseling," and as a clearing house for the consideration of voluntary codes. By July, 1935, a few such codes had been proposed, but doubt as to the legality of any regulations which might be satisfactory to industry acted as a general deterrent.

Selective federal control: Oil. The development of a selective federal control in certain cases directly affecting or involving interstate commerce is indicated by special legislation passed to curb competitive practices in the petroleum and bituminous coal industries. In these industries not only has competition proved itself futile so far as regards achieving stability, but it has resulted in conditions which are of vital concern to the nation as a whole. Petroleum is a wasting asset, and while there may be no immediate danger of a shortage of petroleum products, the time must come when we have to face that situation. Conditions surrounding the exploitation of petroleum lands under private control encourage the rapid exhaustion of the petroleum reservoirs, with the result that production has been at times in excess of the needs of consumption and exploitation has been attended with considerable waste. For many months prior to the passage of the recovery act the price of crude oil was below the cost of production. Since voluntary plans for restricting production had been unsuccessful, and state regulations limiting production and withdrawals from storage had been largely ineffective, it seemed appropriate to Congress to grant to the President special powers, mentioned above in this chapter, for dealing with this industry. In accordance with this authorization the President in July, 1933, issued an order prohibiting the interstate movement of contraband oil and providing penalties for non-compliance; at the same time he established an oil administration under the Secretary of the Interior. This control was bulwarked in August by the setting up of a petroleum code, largely dictated by the government, with the Secretary of the Interior as Code Administrator, which gave the administrator power to fix prices, to establish state quotas of daily production and withdrawals from storage, to regulate imports, and to control the development of new fields or pools. The control was not entirely successful, although it was improved somewhat in 1934 by further regulations

which created a Federal Tender Board to grant tenders for all authorized interstate movements of oil.

In January, 1935, the Supreme Court rendered a decision that the contraband oil regulations issued under authority of the special powers granted in the recovery act were invalid and that this particular section of the act constituted an unconstitutional delegation of legislative power to the President, since it provided no policy or criterion and required no findings of fact to govern the action of the President. To meet this situation new legislation was enacted in February, 1935, declaring it to be the policy of Congress to encourage the conservation of deposits of petroleum in the United States and to protect interstate commerce from the harmful effects of contraband oil, and prohibiting the interstate shipment of contraband oil. The term of this act was limited (as later amended) to June 30, 1939. The President was given power, upon finding of necessity, to require the use of clearance certificates for interstate shipments of petroleum and its products from any particular area and to establish a board or boards to issue these certificates. He was given also authority to prescribe such regulations as might be necessary to carry out the provisions of the act and to require such reports, maps, etc., as might seem desirable. In the event that investigation should show the movement of petroleum to be so limited as to cause maladjustment between demand and supply, the President was given power to suspend the prohibition upon the interstate movement of contraband oil.

An interesting development in the control of the production of petroleum is seen in an interstate compact signed at Dallas, Texas, in February, 1935, effective until August 31, 1939, by representatives of Oklahoma, Texas, California, and New Mexico, and subsequently ratified by the states of New Mexico, Kansas, Oklahoma, Illinois, Colorado, and Texas. Consent to this was granted by Congress in a joint resolution approved on August 27, 1935. The purpose of the compact is to conserve oil and gas, and each ratifying state agrees to enact or retain such legislation as is necessary to attain this objective. It is specifically stated that it is not the purpose of the pact to limit the production of petroleum or to stabilize prices, but to some degree this would seem a reasonable result of

regulations for "conserving gas and oil and preventing the avoidable waste thereof within reasonable limitations."

To assist in the administration of the existing federal petroleum controls and to serve as the contact agency with the Interstate Oil Compact Commission, the Secretary of the Interior established, effective April 1, 1936, the Petroleum Conservation Division. Specifically the Division is authorized to discuss the work of any agency dealing with oil and gas, recommend action on any case brought to its attention, coördinate information, present required data to the Congress, attend oil and gas conferences in which the department is interested, coöperate with the oil-producing states in the study of physical waste and the enactment of uniform oil and gas conservation laws, and maintain contact with other governmental departments whose work deals in any measure with oil and gas. In only one case, the East Texas field, has use been made of the authority to require certificates of clearance for interstate shipments of petroleum. This has reduced the production of contraband oil in that area to a fraction of one per cent of total production.

Bituminous coal: Conditions. Irregularity of production, wasteful methods of mining, potential production normally much in excess of what can be sold at profitable prices, and actual production frequently in excess of this, a laboring force far greater than necessary to man the mines, periodic and bitter labor disputes which have threatened to cut off all shipments of coal — these and many other conditions have been characteristic of the American bituminous coal mining industry for many years.

Because of the importance of conserving the supplies of coal, because of the necessity of assuring an uninterrupted flow of coal, and because competition has demonstrated over a long period of time its inability to bring order out of chaos, there are reasonable grounds for the application of special government measures to the coal mining industry. Under the recovery act a code for the bituminous coal industry was set up, after a long period of acrimonious debate between the operators and the miners. It did not function very effectively, because of the inability of the industry to maintain its established price structure, and of course after the decision in the Schechter case it went the way of other codes. To

take its place in providing control for the industry the so-called Guffey-Snyder coal act was, despite many misgivings as to its constitutionality, passed by Congress and approved by the President on August 30, 1935.

The Guffey-Snyder act. This act declared that the bituminous coal mining industry is "affected with a national public interest" and that the general welfare of the nation requires that it be regulated as provided in the act. To carry out the purposes of the act there was established in the Department of the Interior a National Bituminous Coal Commission of five men appointed by the President. Also established in the Department of the Interior was the office of the Consumers' Counsel, who was charged with the duty of protecting the interests of the consuming public in any proceeding before the commission. A Bituminous Coal Labor Board of three members appointed by the President also was authorized to see that the labor provisions of the act were carried out, to adjudicate disputes, and to act as mediator in any dispute between a producer and its employees "when such dispute is not determined by the tribunal set up in a bona fide collective contract."

On all coal produced in the United States the act levied a tax to be paid to the United States of fifteen per cent of the sale price at the mine, but it authorized the remission of ninety per cent of this tax to producers who filed with the commission an acceptance of the bituminous coal code provided in the act and lived up to it. This code was to be formulated by the commission, but it was required to contain certain provisions specified in the act, among which the following are the most important.

In the first place, the code was required to provide for the organization of twenty-three boards of coal producers, each board representative of a particular geographical area and consisting of not more than seventeen members nor less than three. Within any producing field or fields voluntary marketing agencies might be established under regulations established by these district boards and approved by the commission. These regulations had to pay due regard to the conditions of fair competition embodied in the code in the form of a prohibition of twelve practices specified as constituting "unfair" competition. The district boards were also

authorized, on their own motion or at the direction of the commission, to determine from time to time minimum prices for coal destined for use in the domestic market, which was defined as including the continental United States, Canada, and Cuba. In setting minimum prices the board was empowered to make such classification of coals and price variations as to mines and consuming market areas as it might deem necessary and proper, but it was also directed to fix prices at a figure which would "sustain the stabilization of wages, working conditions, and maximum hours of labor." The commission itself was granted the power to establish maximum prices where it thought these were demanded in the interests of consumers.

The act granted to labor the right of collective bargaining "free from interference, restraint, or coercion of employers, or their agents" and specified that no employee should be required as a condition of employment to join any company union. It was further provided that, when a contract covering maximum hours of work was negotiated between the producers of more than two-thirds of the annual national tonnage and the representatives of more than one-half of the mine workers employed, these hours of work should be accepted by all code members. Wage contracts negotiated in the like manner in any district or group of districts were to be binding upon all the producers of that area.

The day after the bituminous coal conservation act of 1935 came into effect suits were begun challenging its constitutionality, and on May 18, 1936, the Supreme Court, by a five-to-four decision, held the law unconstitutional. Justice Sutherland's majority decision concentrated largely on the labor provisions of the act and found that these were faulty on two grounds: (1) that labor was concerned with production, not commerce, and was therefore not within the jurisdiction of the federal government, and (2) that the labor provisions attempted to effect an unconstitutional delegation of legislative authority in respect to the establishment of maximum hours and wages. Despite a separability clause in the act, the price-fixing provisions were held inseparable from the wage provisions, and the 15 per cent excise tax was said to be, not a real tax, but a penalty.

The Guffey-Vinson act. Following the Supreme Court's decision, Congress reënacted the bituminous coal act of 1935 with such changes as were necessary to meet the criticisms of the Court, and the new Guffey-Vinson bituminous coal bill became law on April 26, 1937. The same administrative machinery, the National Bituminous Coal Commission and the Consumers' Counsel, is established, the only change being an increase in the membership of the commission to seven. A similar organization of district boards is also created.

The tax provisions are more carefully drawn. A tax of one cent per ton is imposed on the sale or disposal of all bituminous coal. In addition an excise tax of 19½ per cent on the sale price at the mine is levied, with a provision for waiving this tax for any producer during his continuance as a member under the Bituminous Coal Code. A change in the provisions for price fixing lodges greater responsibility with the commission, presumably to avoid the criticism of an undue delegation of governmental powers, so that the function of the district board is now limited to proposing minimum prices, which may be accepted, rejected, or modified by the commission. The same principles of price determination are carried over from the 1935 law. A curious reminder of public utility regulation is found in the provision that the commission, if it finds it necessary for the protection of consumers to establish maximum prices, shall fix no maximum price "for any mine which shall not yield a fair return on the fair value of the property." The 1937 law follows that of 1935 by including an elaborate catalogue of unfair methods of competition, which are henceforth forbidden. The Guffey-Vinson act is carefully drafted to apply only to interstate commerce. However it is provided that transactions in intrastate commerce which affect interstate commerce may be brought under the regulations of the code if the commission finds that this is necessary to effectuate the purposes of the law.

The most notable change in the Guffey-Vinson act is the omission of the wages and hours provisions of the 1935 law which the Supreme Court held unconstitutional. Such matters are dealt with under the federal wages and hours law. The labor provisions of the present bituminous coal act are limited to declaring it to be the

public policy of the United States that employees shall have the right to organize and bargain collectively without interference from the employers.

The Coal Commission. On November 30, 1937, the National Bituminous Coal Commission issued its first order, fixing minimum soft coal prices for some 80 per cent of the mines east of the Mississippi and in Iowa. These prices represented increases over those proposed by the operators for industrial and railroad coals, the commission holding that such prices had been depressed below cost of production by competitive conditions. On the other hand, the commission fixed prices for coals for domestic consumption which were lower than proposed by the industry. At the same time certain marketing regulations were promulgated, covering respectively definitions, sales agents, registration of wholesalers, farmers' coöperative organizations, discounts and allowances, limitations of orders, agreements and quotations, spot orders, use of coal analyses, terms of payment, crushing and pulverizing of coal, and "miscellaneous." On January 3, 1938, minimum prices for bituminous coals were fixed for the areas west of the Mississippi. Following court injunctions against the prices fixed for the railroads, the city of Cleveland, and the Associated Industries, Inc., of New York, the commission revoked as of February 25, 1938, all of the 30,000 minimum prices and regulations fixed by its previous orders. The commission then proceeded to hearings for the determination of new minimum prices and regulations, but at present (March, 1939) these had not been concluded.

Conclusions. The conclusion has already been expressed at an earlier point in this chapter that the labor provisions of the NRA, when taken in conjunction with the monopolistic tendencies inherent in code control, could not have been expected to aid in recovery. Nevertheless we have the fact confronting us that during this period of code control there was an increase in business activity. By the end of 1934 factory employment had increased by practically one-third over the level of March, 1933, factory payrolls by more than two-thirds, and the total purchasing power of factory payrolls by one-half, though all remained substantially below the levels of 1929. The average weekly real earnings of persons employed in

manufacturing and mining was slightly higher, though still far below the corresponding level of 1929. The average employee in occupations in the fields of public utilities and of wholesale and retail trade lost slightly on his weekly real earnings between 1933 and 1934, but he had gained so much prior to 1933 that he was still ahead of his 1929 position. In the period between March, 1933, and the end of 1934 the volume of manufacturing production had risen by over 50 per cent, and the value of manufacturing production by over 80 per cent, but they were then still at levels more nearly comparable to 1921 than to 1929. Meanwhile department store sales rose by 35 per cent in dollar value. Activity in the construction industry, under a more direct influence of public expenditures, more than doubled in these months, but the level reached was still very low in comparison with earlier years. In the same period the movement of investment funds into domestic corporate securities issued for new capital increased tenfold but amounted then to only one-seventh of its average volume in the years 1923 to 1925. The year 1932 definitely set a low point for average earnings of business; in 1933 there were, with few exceptions among broad groups of enterprise, considerable reductions in losses, and in 1934 further improvement was shown. There was a considerable increase in the aggregate dividends distributed by a large number of corporations in 1934. During the fall of 1935, after the decease of the NRA the upward tendency in business activity was continuing.

The quantitative effect of the NRA labor program on business recovery is exceedingly difficult, if indeed not impossible, to measure. A number of other factors were working on prices and the market at the same time — government spending on a huge scale, attempts to raise prices through currency manipulation, as well as certain natural forces of recovery. For reasons which have already been advanced, it is impossible to give to the labor program of the NRA credit for much if any of the improvement in business conditions. In fact it is probable — and this is the considered view of many competent critics of the program — that the effect on the whole was to retard the process of recovery, chiefly by discouraging private industry. Without doubt a part of the gain in business activity should properly be ascribed to the continuance of large-

scale spending of borrowed funds by the government, but it should be realized that the program itself has had a discouraging effect on many lines of private industry. As previously stated, there is reason to believe that an increase in consumer spending brought about by such governmental activity can have no more than a temporary influence on business recovery. In the total flow of purchasing power intermediate spending by business enterprises is of far greater importance than consumer spending. Until some approach to a reëstablishment of balance among economic factors becomes apparent and business confidence in the prospects for continuing profits is consequently created, business spending and business investment are discouraged and active business recovery is delayed. The proceeds of any new consumer spending tend simply to lie stagnant in idle cash balances of business enterprises or in idle excess reserves of banks.[1]

The authoritative maintenance of such adjustments between industrial costs and prices as to produce an increased purchasing power of wages and presumably some improvement in profits presents a difficult enough problem of control, but the administration in 1933–1935 was attempting to do three other things at the same time. It was attempting, first of all, to increase the prices of farm products more rapidly than those of industrial products; in the second place, it sought to protect the consumer against excessive increase in the prices of goods produced by other economic groups than his own; and in the third place, it was attempting to secure a rise in the general level of prices. In short an enormously difficult and intricate excursion in economic planning was embarked upon. Even if such a venture be thought desirable, it is doubtful if we yet have a sufficiently detailed and exact knowledge of the interrelated workings of economic processes to make the attempt feasible.

The effective elimination of unfair competition throughout business in general is an objective worthy of the most careful consideration, but it is not to be achieved in a few brief months by

[1] The reader who wishes to pursue this topic further is recommended to consult Leverett S. Lyon and others, *The National Recovery Administration*, Brookings Institution, 1935.

any statute, code, or administrative ruling, whether imposed or voluntary. It is conceivable that closer business coördination, fostered by working agreements in industry, may develop a more effective business opinion against unfair competitive methods and thus diminish their violence, but any program looking toward this must face the danger, on the one hand, of encouraging monopolistic practices if too great powers of self-government are given to industry, and on the other hand, of reducing the play of the forces of individual initiative if an active policy of government control is to be pursued. To preserve competition while exercising control over its destructive aspects is most difficult, since to the extent that control becomes successful the operating flexibility of competition tends to be reduced. In industries containing hundreds of small enterprises control must become so detailed as to leave little room for free initiative, and even then the minute legal stipulations will scarcely conquer forthwith the undesirable business practices of years.

XLVII

GOVERNMENT IN INDUSTRY

The motives for government industry. The activities of government fall naturally into two groups: (1) those associated with carrying out the primary functions of government, such as providing for national defense, justice, security, education, and the like, and (2) industrial activities. When government enters the industrial field, it is generally either because its control of some particular enterprise is considered essential to the performance of one or more of its primary functions, or because it is felt that some essential commodity or service can thus be provided for the people more economically or more equitably than by private industry.

When the telegraph and telephone were first introduced, it was recognized by the governments of the leading nations of Europe that these instrumentalities were destined to be of great military importance, and it was felt that it would not be safe to leave them in private hands. For this reason the telegraph and telephone have from the beginning been government enterprises in most European nations. This consideration is also the chief explanation of the prevalence of government ownership of railroads in Europe. The United States, remote from the militaristic atmosphere of Europe, has not been moved by this motive, but it is interesting to note that at the time of the World War military considerations led the United States to assume temporary control of the railroads and the telegraph and telephone services.

The obligation of the government to safeguard the national resources vital to all industry leads to the holding of a certain domain and the assumption of certain industrial undertakings under the general program of "conservation." Thus public forests are a common field of government ownership, and public mines are occasionally found. Protection against disease sometimes requires city governments to take a hand in such enterprises as milk stations, bath houses, housing projects, etc.

Certain industrial enterprises, not so closely associated with particular governmental functions, are regarded as essential in a general way to the functioning of government. An obvious case in the United States is the government printing office. The most important and universal example is the post office. In most European countries the postal service was originally undertaken by the government for military or other direct governmental purposes. But the principal reason for the government post office today is the social and political motive. Easy and cheap communication, open to all, tends to bind the people together with common ideas and ideals — to foster a "national spirit." Public opinion has opportunity to crystallize and to express itself, and the government may be informed of the wishes of the people. These ends are felt to be vital to the well-being of the modern democratic state, and it is feared that they would not be so fully accomplished if postal service were left in private hands.

Outside the realm of its primary functions, government may accept the obligation to assure the people a supply of essential products or services either (1) because the enterprise is too vast or uncertain to invite private capital, as in the case of the Panama Canal, the government railroads of Australia, India, etc., or (2) because the return from private enterprise is too remote or speculative, as in the case of forestry, or (3) because of the desire to protect the consumers from the exactions of private monopoly. In addition there are occasions when the fiscal motive leads the government to take over the manufacture or sale of some product for the sake of the monopoly profit, as in the tobacco monopolies of France and Italy and an occasional salt monopoly.

The public domain: General historical trend. Originally the entire national territory was in most countries the property of the state or of the sovereign; such proprietorship was the basis of the feudal system. The trend in modern time has been almost universally in the direction of alienation of government lands into private ownership, and few important nations today retain any extensive landed domains.

The United States has, in a remarkably short time, passed to private ownership a domain of vast extent and fabulous wealth.

Today the United States retains ownership, outside the forest reserves, of very little land that is not arid or rocky or otherwise unsuited for cultivation. Absence of large land holding also generally characterizes the American states and municipalities. The government does of course need a certain amount of land for the performance of its primary functions: sites for public buildings, military reservations and forts, agricultural experiment stations, public parks, etc.; these obviously present no exception to the above historical generalization.

Here is not the place to enter into lengthy discussion of this interesting development; suffice it to note that, in harmony with general principles with which the reader is now familiar, experience has proved that the most efficient use of the land comes as the result of private ownership and operation.

Forest reserves. There is however one real exception to the general conclusion, theoretical and historical, as to public ownership of land; namely, the forests. It is known that the presence of a forest cover on the mountains and hills is essential to protect the sources of the streams and maintain an even flow of water to the low lands; otherwise there are floods in the spring and after heavy rains and dry streams at other times, involving irreparable losses. Since the private owners of forest lands have ordinarily no inducement to preserve their forests for this purpose, the government is called upon to take the responsibility in the interest of the whole people. On the other hand, realizing the vital importance of lumber and other wood products and foreseeing the decline of the natural sources of supply under private ownership, most national governments have accepted responsibility for maintaining a supply of forest products for future generations. Large and increasing public forest reserves are thus to be found among the nations of the world, most of whom have in this case reversed their policy of alienation and adopted a program of increasing the state-owned domain. Forests reserves are held, not only by national governments, but also by states and departments and even by municipalities, notably in Germany, Switzerland, and France.

In the United States, the federal government has for a generation or more followed the policy of forest conservation. Federal lands,

together with Indian lands under federal control, include approximately 100 million acres of commercial forest area. There are also about 11 million acres of such forest owned by states, counties, and municipalities, so that the total area of publicly owned or controlled commercial forest stands at 111 million acres, or more than one-fifth of the 495 million acres of commercial forest land in the United States (all of these figures excluding Alaska). The total area in public forests, including grazing, barren, and brush lands, is much greater than the area classified as in commercial forest, the national forests alone containing 175 million acres, including 21 million acres in Alaska.[1] The public forest area is being increased from year to year by purchase, by exchange of private lands for government timber, and by gift.

In Europe the density of population, the prices of forest products, and the high stage of scientific forest management have made it possible for the public forests, with certain exceptions, generally to show a financial profit. In America this has not generally been possible, though the United States forest administration is approaching a balance between cost and revenue. The success of the public forests is of course not to be measured solely by the financial results. Those forests which must be held for protection of stream flow would be socially profitable even though, with economical management, they showed a financial loss.

State railways. Government owned railroads are a commonplace in most parts of the world. It is estimated that 42 per cent of the world mileage is state owned and operated.[2] In Germany the railway system was, prior to 1920, owned by the several states. These railways were on the whole efficiently managed and financially successful, in 1912 they had an invested capital of about eighteen billion marks, of which about twelve billions were represented by public indebtedness. It was estimated that their excess of revenues over operating expenses showed a return of something over four per cent upon the invested capital. After the World War the railways were taken over by the imperial government. The Reich owns all the common stock of the German Rail-

[1] Forest area data furnished by the United States Forest Service.
[2] P. H. Middletown, *Railways of Thirty Nations*, 1937, p. 291.

way Corporation, organized in 1924, which owns and operates ninety per cent of all rail trackage. The Reich has full control of all railroad regulation. The imperial government's operations have not enjoyed the financial success which was formerly attained by the state railways. There were deficits in 1932 and 1933, and a slight excess of receipts over expenditures in 1934 and in 1935.

In France the railroads were until recently privately owned and operated, with the exception of one government line. In 1937 a decree was issued setting up a National Railway Company to absorb all the French railroads on January 1, 1938. Fifty-one per cent of the stock of the company is held by the state, and the remaining 49 per cent by the companies. Management of all railroads by the National Railway Company will continue until 1982, when the charters of the several private companies expire and the assets of these railroads become the property of the nation. Financially the French railroad business, whether public or private, has not been especially successful. In order to force development of the private lines, the government has guaranteed interest and has had to make good considerable deficits of the private companies. The cumulative deficit on this account amounted to 27 billion francs up to 1937. The one government line has generally not succeeded in earning the interest on its capital, even during the decade before the World War, when some of the private lines were making profits.

The financial results of the Swiss state railways have of late been unsatisfactory. The state railways of Belgium were not profitable, and in 1926 they were leased to a private company.

In such countries as Australia, Canada, and the United States the railroad problem has been quite different from that with which the older, more densely populated countries of Europe have had to deal. In Australia some of the state railways have produced surpluses, some deficits, after allowing for interest; the deficits on the whole exceeding the surpluses. In Canada, where there is both private and public ownership of railways, the Canadian National Railway has been operated at a very heavy loss. On the other hand, the railways in these countries have served to aid rapid settlement and development of the national resources. The same

purpose has been accomplished in the United States mainly through the agency of private enterprise, although with substantial public aid in the form of land grants, stock subscriptions, and guarantee of bonds. The early experiments in state railroad enterprise in the United States were notorious failures. On the other hand, the policy of state ownership has been followed in the case of the Alaskan government railroads.

During the World War the entire national transportation system was taken over by the United States government and held from December 28, 1917, to March 1, 1920. This action was necessitated by the inability of the private companies to respond adequately to the extraordinary demands of war. The government administration introduced many changes which facilitated efficient operation for war purpose, consolidations and rate increases previously forbidden to the railroads were introduced, and wages were advanced. The government operation was conducted at greatly increased cost, producing a total deficit estimated at more than a billion and three quarters dollars.

State telegraph and telephone service. These services are commonly owned by government in Europe, and the record is on the whole similar to that of the railways. The British telegraph system, taken over by the nation in 1870, has operated continuously at a heavy financial loss, though this has been materially decreased in recent years. The state telephone service has been efficient and financially successful in some countries, as in Great Britain, Germany, and Switzerland, while unsatisfactory in others, as in France. The United States has had virtually no experience with government operation of these services. It is of course true that the United States undertook the operation of the telegraph and telephone services for military purposes for one year, beginning August 1, 1918, but the period of control was too short and the objectives of control too remote from those of ordinary peace times to shed light on either the virtues or the defects of government industry.

The Panama Canal. Limits of space forbid any general survey of state enterprise in canal transportation; we shall have to be content with brief reference to a single example. The Panama Canal, doubtless the greatest engineering feat of all time, was

built under complete government control after private enterprise had proved inadequate to the task. A French company, the Universal Interocean Canal Company, under the direction of Count Ferdinand de Lesseps, the builder of the Suez Canal, worked for eight years on the project only to end in failure. When the United States determined to carry through the undertaking, it was contemplated that the actual work would be carried on by private companies on a contract basis. The attempt to let the work out by contract was not successful, and Colonel George W. Goethals, the engineer in charge, was forced to resort to direct employment of the entire staff.

The canal was dug primarily to serve the interest of American defense and American commerce. Since the opening in 1914 there has been a steady increase in the tonnage carried, so much so that there has been some talk in responsible quarters of the necessity of building another canal across the Isthmus in the near future. As an aid to commerce the canal has been a success. Financially the verdict is not so favorable The accounting system employed is such as to render impossible an accurate statement of the invested capital upon which interest and amortization should be earned. The best estimates that can be made appear to indicate that up to 1930 there had been accumulated by the canal and associated enterprises a net deficit of about 143 million dollars. The deficit of the canal proper was considerably greater than this, being offset by surpluses from the operations of the Panama Canal Railroad and the Panama Canal Steamship Company. During the last few years of the decade from 1920 to 1930 the business was financially successful, earning about 3 per cent on the investment, which was being applied to reduction of the deficit.[1] The net revenues of the canal in the fiscal year 1936, combining canal operations proper and business operations, were 15½ million dollars. In 1937 the combined net revenue was 14 millions. The capital investment at the beginning of the year 1937 was $549,874,827. The net revenue of that year represented a ratio of 2.56 per cent on the investment, as compared with 2.83 per cent in 1936.[2]

[1] Mills and Starr, *Readings in Public Finance and Taxation*, 1932, pp. 175–195.
[2] *Annual Report of the Governor of the Panama Canal*, 1937, p. 3.

The Panama Railroad Company, a corporation acquired by the United States government in 1904 at a cost of $7,000,000, has had a good record for efficiency of operation, which is reflected in earnings and in dividend payments to the United States Treasury. From 1914 to 1933 dividends have amounted to $8,744,905, or at an average rate of 6.2 per cent on the capital stock.[1] The net profits of the Panama Railroad Company were $1,077,987 in the fiscal year 1936 and $1,358,595 in 1937.[2]

The United States Post Office. The foregoing survey of the field of state industry has been brief and of the most general nature. As a single example for more detailed study, the United States Post Office will be most useful, representing as it does the one enterprise that is everywhere government owned, as well as being the one best adapted to throw light on the problem of national industry in the United States. That the United States Post Office is a great industrial enterprise is sufficiently evidenced by its payroll of 300,000 employees and its annual receipts of over half a billion dollars.

Postal rates. Assuming that the more obvious features of the postal organization are known to everyone, let us proceed at once to a somewhat critical analysis of the scheme of postal rates, which is the most characteristic feature of the business. Mail matter is divided into four classes, as follows:

First class. Letters and all other written matter, as well as all packages that are sealed. *Rate:* 3 cents per ounce; for local delivery, 2 cents; postal cards, 1 cent.

Second class. Newspapers, magazines, and other periodical publications. *Rate:* (a) When mailed by any person other than the publisher or news agent, 1 cent for each 2 ounces. (b) When mailed by the publisher or news agent, that part of the publication which does not consist of advertising matter pays a basic rate of $1\frac{1}{2}$ cents a pound; on advertising matter the basic rates are from 2 cents to 10 cents a pound according to the distance; if sent to an office in the same county that has neither city nor village delivery, there is no charge.

Third class. All other printed matter, such as books, pamphlets,

[1] Dimock, *Government-Operated Enterprises in the Panama Canal Zone*, 1934.
[2] *Annual Report of the Governor of the Panama Canal*, 1937, p. 3.

circulars, etc., and merchandise, not exceeding 8 ounces in weight. *Rate:* $1\frac{1}{2}$ cents for each 2 ounces.[1]

Fourth class (*parcel post*). All matter weighing more than 8 ounces, except first class and second class matter. *Rate:* 7 cents per package and up, depending on weight and distance.

Air mail. The rate is 6 cents per ounce.

The flat rate. This scheme of rates has been presented thus in detail because it raises certain important questions of principle in connection with the postal business. We notice first that in the first and third classes and to a considerable extent also in the second class the rates charged are based upon weight alone without reference to the distance carried. This principle of the flat rate is a characteristic feature of the postal business — at least as regards letters — all over the world. It is justified on two different grounds. From the administrative point of view, by far the greater part of the cost of handling mail matter is caused by collecting and delivering at the two ends regardless of the distance carried. It becomes possible to average the carrying cost without undue arbitrariness and of course with much gain in simplicity.

The other and more fundamental basis of the flat rate charge is the political purpose to make the post office a means of uniting the whole people through cheap and wide-spread communication. While in general the cost of carrying is relatively insignificant, this is not true of the more remote and inaccessible places. Under the flat rate system, residents of those regions are not called upon to pay the very heavy costs of their own postal service but are given the advantage of the low average rate at the expense of the residents of the more thickly settled regions. The additional burden on the latter however, when spread over the whole business, is insignificant and is accepted without complaint in view of the important ends to be served.

The foregoing is true of letter postage. The balance of advantage and disadvantage swings the other way when the post office goes extensively into the carrying of merchandise. Here the carry-

[1] There is a special rate in this class of 1 cent for each 2 ounces on "books catalogs, seeds, cuttings, bulbs, roots, scions, and plants," and reduced rates under certain conditions.

ing cost, depending on distance, is by no means insignificant, and the flat rate cannot be justified. The flat rate was employed in all four classes of the United States postal service until the introduction of the parcel post with the "zone" system on January 1, 1913. It was not till 1917 that the principle of differentiation was extended to periodical advertising matter in the second class.

Rate discrimination. The most striking feature of the United States postal rate structure is the wide differentiation between the rates of the several classes of mail matter. Reduced to a pound basis, the primary rates of the first three classes are respectively 48 cents, 1½ cents (publishers' rate), and 12 cents (books 8 cents). Letters are charged thirty-two times the rate imposed upon periodicals, six times the rate on books, and four times the rate on small merchandise packages; books are charged more than five times as much as periodicals. Of course the high pound rate for the first class is in part justified by the two facts that first class mail is accorded a preferred service and the individual packages are so much smaller than in the other classes. Yet if we reduce the rate to a piece basis, the discrepancy is still very great. The average revenue per piece of mail matter in the fiscal year 1937, as estimated by the Post Office Department, was as follows. first class, 2.82 cents; second class, .54 cents; third class, 1.34 cents.[1] The first class charge, even on this basis, is more than five times that of the second class, in spite of the relatively small size of the first class pieces.

The fiscal aspects of this postal rate structure are significant. The United States Post Office does not have a system of cost accounting such as is considered essential in any up-to-date private business. However the national interest in this problem is such that several special investigations have been made; the last, made by the Post Office Department and completed in 1924,[2] presents an analysis of revenues and expenditures for the year 1923. Since 1925 the Post Office Department has made annual estimates based on sample tests in certain post offices. The following figures,

[1] Postmaster-General, *Report, 1937*, p. 113.
[2] *Cost of Handling Mail Matter*, 68th Congress, 2d Session, Senate Document No. 162.

which agree closely with the results of the investigation of 1924, are selected as pertinent to our present inquiry.[1]

ALLOCATION OF REVENUES AND EXPENDITURES, UNITED STATES POST OFFICE
FISCAL YEAR 1937
(in millions)

Class of service	*Revenues*	*Expenditures*	*Loss*	*Gain*
Paid first class	$397	$273	—	$123
Second class	24	113	$ 89	—
Third class	72	92	20	—
Fourth class	132	146	14	—
Foreign	17	32	14	—
Penalty, franked, and free	–	15	15	—
Special services	67	84	17	—
Total (mail and special)	$710	$756	$169	$123
Net loss*			46	

* Certain unassignable and unrelated items are disregarded.

The first class business shows a profit of 45 per cent of the cost, the result of collecting $1.08 a pound or 2.8 cents a piece for a service which costs only 74 cents a pound or 1.9 cents a piece. On the other three classes there is a loss, moderate on the fourth class, considerable on the third, and very heavy on the second class. Here the cost is four and a half times the revenue; the post office collects 1⅞ cents per pound for a service which costs it 8⅓ cents to perform.

The government has a monopoly of the business of carrying first class mail, under the protection of which it is able to charge high rates and make a large profit. But the profit is wiped out by the loss on the second class. The result is a subsidy to the periodical publishers at the expense first of the users of first class mail service and second of the general body of taxpayers, who have to make good the deficit on the postal business as a whole. This differentiation has many interesting aspects, political and social as well as economic, into which we cannot enter. For example it means a large subsidy to periodical advertising and a heavy discrimination against books in favor of newspapers and magazines.

The postal deficit. Throughout the world in general the postal business is operated, not for a profit, but with the idea of making the revenues about cover the cost. In the United States the post

[1] Postmaster-General, *Report, 1937*, p. 111.

office has always been operated at a loss, and the loss is actually considerably greater than the official figures show on account of the absence of charges for interest, depreciation, and other essential items in the accounts. The annual deficits must of course be met out of the general funds raised by taxation. It is not likely that the second, third, and fourth classes could ever be made self-supporting, whatever the rates, so long as they are in competition with private enterprise. But it would be possible so to adjust the rates in each class as to produce for all services except the first class a moderate deficit only, which could be balanced by a moderate profit on the first class.

The general verdict of those who have examined the conduct of the American postal business is that, while it accomplishes its main purpose, it does so in an unnecessarily arbitrary manner, its subsidy to periodical advertising being especially difficult to justify; it operates without any real accounting system and with a very large and unnecessary deficit; and these results are chargeable, not only to legislation by Congress, but also to serious lack of sound business judgment in the management of the enterprise.

Industrial ventures of the American states. In the United States the states are at present generally holding aloof from industrial enterprise. In the first half of the nineteenth century several of the states ventured rather extensively into the industrial field, particularly in railroading and banking. These state enterprises were almost without exception disastrous failures; the states lost heavily and in certain cases defaulted upon the bonds by which the capital investment had been obtained. The lesson was so thoroughly learned that to this day the states have been generally quite willing to keep out of industry. There are, it is true, certain states owning grain elevators, printing plants, etc. The most important example is the State of North Dakota, which, under the lead of the ill-fated Non-Partisan League, set forth in 1919 upon a program of state banking, flour milling, and other enterprises. The bank had a capital of $2,000,000, borrowed by the state; after losing about a million and a half in the crisis of 1920 and the succeeding years of depression, it has lately confined itself largely to handling the state's business, including a farm loan enterprise.

The state flour mill and grain elevator were likewise unfortunate financially, losing by the end of 1927 about one and a half million dollars on a capital of four and a half millions borrowed by the state.[1]

Extent of municipal industries. The field of municipal industry has been widely cultivated in European countries. Municipal undertakings are numerous in water, gas, electricity, tramways, public markets, public baths, and the like. The United States, where public opinion has in general been more apprehensive with regard to the efficiency of public enterprise and more convinced of the merits of private business, there has been less readiness on the part of any grade of government to engage in industrial enterprises. Nevertheless there has been some development in this direction by the municipalities. The picture of the accomplishments is not entirely clear, inasmuch as our chief source of information, the Bureau of the Census, covers in its principal annual reports only the cities with population exceeding 30,000, but a summary of the more significant facts may be of interest.

The water supply is commonly under the direct control of the city and town, and the trend is clearly in the direction of an increase in public ownership. In 1927, of the 25 cities with a population of 300,000 or more, all except one (Indianapolis) had public water-supply systems. In 1929 in communities of 5,000 to 10,000 population 75 per cent were so operated, while in the cities and villages below 5,000 in population 85 per cent of the plants were publicly owned and operated. Another indication of the importance of public water works is the fact that in 1929 2,692 plants out of a total of 3,478, or 77.4 per cent, were publicly owned.

The number of municipally owned electric light and power establishments[2] showed a fairly rapid increase up to 1923. From that date on the number both of municipal and of commercial establishments has decreased, very largely as a result of consolidations. As is natural, the number of private establishments decreased more rapidly than was the case with municipal plants

[1] H. L. Lutz, *Public Finance*, 1929, pp 198–199

[2] According to the Census terminology an "establishment" is a plant or plants under single ownership.

because of the greater possibility of securing economies by means of the erection of large central generating plants, the building of high-tension transmission systems, and the abandonment of a number of small generating plants. The result is that there are actually fewer commercial electric light and power establishments in the United States today than there were in 1902. According to the Census reports there were 2,805 such establishments in 1902 and only 1,627 in 1932, whereas there were 815 municipal plants in 1902 and 1,802 in 1932. But by no other standard than numbers can the public plants be considered more important than the private establishments. In 1932 investment in plant and equipment by private companies amounted to over twelve billion dollars, while the municipal establishments had only about 540 millions; the net income of the private companies was reported as over 538 million dollars as compared to 37 millions for the municipal plants; and the municipal plants accounted for only about five per cent of the total output of electrical energy.

Technically the large private establishment, with huge generators well located and able to serve a number of localities through a high-tension transmission system, has had such an advantage over the municipality with a small steam-generator serving only the immediate community that in recent years a number of public plants have been turned over to private ownership (834 from 1924 to 1927), and in many other cases the municipalities have closed down their own generating plants, finding it cheaper to purchase power from private companies for subsequent distribution. The fairly recent development of the Diesel engine, using low-grade fuel and having a flexibility in operation which apparently makes it peculiarly adaptable to power production, has robbed the large generating companies of some of their advantages in production and has slowed down the process of closing out high-cost municipal plants.

In the gas industry municipal ownership has not made much headway, and the trend for the last twenty years has been toward a decrease in the number of municipal plants. In 1929 there were only 66 municipal plants out of a total number of 1,614,[1] and these

[1] W. E Mosher and F G. Crawford, *Public Utility Regulation*, 1933, p. 733.

plants accounted for little over one per cent of the total production of gas. The development of street railways is quite insignificant. Public fuel yards, piers and terminals, public baths, markets, etc., add somewhat to the total of municipal industry, though these additions are not capable of measurement.

Financial results of government industry. As regards the financial operations of municipally owned public utilities the evidence is incomplete, conflicting, and often unreliable. Frequently the accounting systems of municipal enterprises are not adequate or reliable; the Census Bureau complains regularly of the untrustworthy accounts upon which it has to base its statistics. Apparent net operating revenue must in many cases be reduced because of failure to make proper allowances for interest, depreciation, or amortization. Services, such as street lighting, may be provided free to a city without an appropriate charge on the books, and similarly many services may be furnished to the utility by other departments of the city without appearing on the accounts. In the face of these facts any sweeping generalities would be unwarranted and unsound. On the whole the financial record of public industry is unfavorable; there are conspicuous cases of financial success, but the number and magnitude of the industries operated at heavy deficits are more impressive; and in most cases, in America at any rate, the actual facts are far less favorable than the published figures on account of the prevalence of faulty accounting methods.

The question of public industry is not however to be settled on the basis of the financial results alone. When deficits are the result of inefficiency, their presence weighs in the balance against public industry. It is quite possible however to show a deficit with efficient operation, simply because that was the purpose of the public industry. This means that the people choose to have a certain service rendered for the consumers at less than cost. Private industry will not do this; not at any rate without a subsidy. Therefore the government must take the responsibility, and the taxpayers must make up the difference. Such a decision is within the province of any body politic; the question turns upon general considerations of social policy. It will for example be generally granted that

protection forests should be maintained by government even if the service could not be made to pay its cost. The Panama Canal might have been justified on grounds of public policy, even though it did not succeed in paying for itself. Canada and Australia may consider that the net results of their rapid railway development outweigh the financial deficits of the government railroads. One must only remember that such conclusion is, in each case, to be demonstrated, not taken for granted, and that the burden of proof is fairly placed on the advocate of public ownership. This conclusion furthermore does not render negligible either efficient management or honest accounting. If industry is to be run at less than cost, it becomes if possible all the more important to make the cost as small as may be and to know what the real cost is.

Cost of service. Comparison of commercial and public enterprises as to service and rates, which would be the more significant criteria, is rendered inconclusive by the character of the data as already described. The desirability of securing more reliable information regarding the possibilities, financial and otherwise, of governmentally owned electric enterprises was one of the motives which led the United States government in 1933 and thereafter to approve such projects as the Tennessee Valley Authority and the Columbia River developments. These involve, among other things, the generation and distribution of light and power on a huge scale and the stimulation of a demand for electricity by low rates and by an intensive educational campaign to promote the use of electrical appliances. The sponsors of these projects promise us examples of government industry at its best — the strictest accounting, the most reliable and progressive management, and of course a scale of operation sufficient to secure the maximum possible economies. It is claimed therefore that they will serve as "yardsticks" by which to measure the accomplishments and check the claims of privately owned utilities. Of the interest and significance of these experiments there can be no doubt, provided the performance is as fair as the promise. Serious question has already arisen regarding the cost accounting of the TVA. The Congressional investigation of this agency undertaken in 1938 may give material for more mature conclusions than the present state of our knowledge warrants.

The problem of government industry: Its advocates. Those who favor government industry may be classed in two main groups: those who advocate the complete nationalization of all industry, and those who believe that within the framework of a system of private capitalism the government can profitably undertake certain types of industrial enterprises. In the first group we find the irreconcilable critics of the present order who are convinced that the weaknesses of private capitalism are so fundamental that the only hope of stable prosperity lies in the direction of socialism. With their program we are not concerned at this point, our task is rather to present the point of view of those who believe in the present economic system, but who feel that it can best be preserved by removing from private hands those parts of it which have become monopolistic in character or honeycombed with abuses.

The merits of government ownership. In determining the issue of public *versus* private industry, recourse will inevitably be had to general principles as well as to statistical evidence. Sentiment in favor of government industry may be aroused by contemplation of the problems of industrial combinations and monopolies and the problems of the railroads and other public service enterprises. The argument in favor of government ownership lays stress upon the discriminations and unfair practices which accompany private competitive enterprise, and it becomes stronger still when private monopoly arises.

Attention is directed also to the large sums paid for interest on private capital and the large profits made by private enterprise, and it is claimed that these might as well accrue to the whole people in the form of lower rates. Furthermore the claim is often made that public enterprise is apt to furnish better service to the consumer, since its sole *raison d'être* is service and the question of profits does not have to be taken into consideration.

Interest and profits. It is evident at the outset that the claims of the advocates of government ownership cannot be substantiated with respect to the saving in interest. This is a charge which must be met whether the industry be public or private. It is a payment essential to the accumulation of capital, and so long as saving is left to private initiative the government cannot get capital

without the payment of interest any more than a private entrepreneur. If attention be called to the fact that government bonds bear a lower rate of interest than industrial investments, the answer is that the financial difficulties which are the heritage of the World War have already deprived many governments of this advantage, and the whole advantage would tend to disappear in proportion as government became involved in the risks of industry.[1] If the taxing power be relied upon to make good the losses of unsuccessful government enterprises, this is merely robbing Peter to pay Paul.

The same general considerations apply to the alleged saving of profits. Profits, as we know, are paid for the service of risk bearing, and we have concluded that the payment, considering losses and gains together and taking industry as a whole, is distinctly moderate. If the risks of industry were transferred from the shoulders of private entrepreneurs to the government, they would still have to be paid for. Governments cannot take the profits and avoid the losses. The government industry would have to charge enough to enable it to make experiments, assume risks, and stand the inevitable losses. There is no reason to believe that the cost of risk bearing would be any less if undertaken by the government, unless it were to avoid risk by permitting industry to stagnate.

Profits and local public utilities. While the above considerations apply to industry in general, it is possible that in the special fields into which the municipalities particularly have ventured there may be a saving to the consumer by reason of a reduction in the volume of profits. The purveying of water or electricity is generally a monopoly, whether the plant be publicly or privately owned. As compared with competitive manufacturing for example the risk is slight. With private ownership pressure is constant to maintain existing rates or secure higher ones. Quite commonly a rate which at one time yielded reasonable profits has, with a thickening of the population, given enormous profits. Under private management the action of the public service commission in reducing rates is slow and halting; under public ownership, assuming for the moment that public management be as efficient as private or even slightly

[1] The very low interest rates on United States government obligations that prevailed during the depression following the crisis of 1929 were due to special causes.

less efficient, either the consumers would get the benefit at once in the form of lower rates or the taxpayers would benefit from lower taxes.

Corporate abuses. In their advocacy of government industry, some lay greater stress on the necessity of eliminating certain undesirable aspects of corporate control than on the actual savings which government industry might secure. In other words government ownership is to them a necessary substitute for government regulation. They point out that the problem of watered stock, which frequently affects the rate base and therefore the cost of service to consumers, would be done away with, that the looting of utilities by insiders would be impossible, and that speculation in securities of utilities would be no longer possible.

But more fundamental is their belief that regulation of public utility rates has been halting and ineffective and that it is idle to place reliance upon regulatory commissions to secure rates which are only reasonably remunerative. It is not within our province to discuss the problem of regulation of public utilities, but it is perhaps not out of place to note that in many states appointments to regulatory commissions have been on a political basis, that the task of proving that a given rate is unwarranted involves long-drawn-out and very expensive litigation, and that the problem of regulation has been exceedingly complicated because of the interstate character of many of our holding companies and the absence of federal regulation. The task furthermore of evaluating the property of the companies for purposes of rate-making is both theoretically and practically a difficult and lengthy process. It is however possible to admit the existence of these conditions, as do most authorities, without reaching the conclusion that government industry is the only solution, for the maximum possible effectiveness of regulation has certainly not been reached.

Politics. There are certain positive weaknesses which seem inherent in government management. In the first place, there is weakness in personnel and discipline. Politics controls, more or less, the selection of employees, both managers and subordinates. There is no disguising the fact that public employees are frequently chosen on grounds of political affiliation or service rather than

because of fitness for the work. Examples are a commonplace in America in all grades of government, and the practice is accepted almost as a matter of course. This means necessarily that the government usually has less able managers and a less efficient labor force. Politics also interferes with discipline; those in charge hesitate to enforce adequate control over their subordinates or to discharge the incompetent, for fear of political protest. For the same reason the managers have often been prevented from installing approved labor-saving machinery, cost accounting systems, or up-to-date business methods. Thus is efficiency sacrificed and industrial progress hindered.

Political sentiment also tends to prevent the payment of salaries sufficient to attract into the public service the most capable managers and technical experts. One of the serious handicaps with which the head of any government service has to contend is the constant drain of his most competent subordinates into private business attracted by the lure of higher pay.

On the other hand, the utilities are frequently in politics and exerting pressure on legislation through lobbies and organized and "educated" public opinion. Society would gain something by removing this danger to democratic government, although it would probably stand to lose more from the development of bureaucracy and the political control of government industry.

The motive in business. Public industry is weak as compared with private with respect to the motive of the managing head. The private business man puts the utmost of his skill and energy into his business because he has a personal stake in it. He has risked his own capital; he faces the opportunity of profits and also the danger of loss; his economic well-being and the comfort of those dependent on him hang upon the success of his enterprise. Even if he be only the hired manager of a corporate business, he finds the same standard held up to him by the stockholders, who look to him to protect the capital and produce profits for them. His position and his fortune depend on making a success of the business under his charge. Self-interest is the spur which drives the private business man to make his business efficient.

To the public official this motive is weak or absent. He has not

risked his own capital in the government industry. If profit made they will not accrue to him, and he will not have to bea financial consequence of failure. The continuance of his pos and his further advancement do not depend upon the judgmen critical body of stockholders as to the efficiency with which h managed the business. His position is more often dependent pleasing the general body of voters or the leaders of his pol party. This approval may be won by easy-going treatment o subordinates and showy service to the public, rather than by hard test of business efficiency.

This is of course not to assert that there never is nor can be cient management in public business. There are capable animated with a keen sense of public service, and there are munities in which the citizens insist on able and honest mar ment. There have been conspicuous cases of high-grade leade impelled by motives of patriotism and public service, especial times of national emergency such as war. On the other hand, considerations do not deny the importance of the differenc motive between the managers of private and public industries

Conclusion. The problem of public *versus* private industry still remain undetermined, in the sense that it is impossible to a at any single broad generalization as to the relative merits o two forms of ownership. It is not at all likely that the prol ever will be solved in the form of such general conclusion. I two countries and with respect to no two classes of industry ar conditions of the problem the same. The decision as to any indu may go one way in England, the opposite way in America; in given country the decision may be for municipal water works against municipal gas works. The defects of private industry compel public ownership in one case, as of the forests on the w sheds; while the weaknesses of public ownership may lea substitution of private industry, as in the case of the Philadel gas works some years ago. In one state effective regulation give satisfactory rates and service, while in another a weak c mission may force municipalities to assume the risks of gov ment industry. It may fairly be said that the mass of evidence the body of general principles bearing on the question are such

place the burden of proof upon the one who advocates government ownership. From this position the public will undoubtedly feel its way along the road of economic progress, determining each case upon its merits in the light of such knowledge as may be at hand.

One general conclusion may perhaps be stated with some confidence. The public is not likely to tolerate unlimited private monopoly in the necessities of life upon any large scale. In those basic lines of industry where monopoly appears inevitable, the choice is between private industry with government regulation and government industry. Attempts at government regulation have thus far been generally short-sighted and bungling. We may now be upon the threshold of the first real test of this program; should it fail, the alternative of government ownership will still be open. In the meantime there may well be no definite conclusion — no decided swing either to public or private ownership -- but rather a balance between the two, with private ownership the fate of the inefficient public plant, and public ownership a threat and a spur to the privately owned public utility.

SUGGESTIONS FOR FURTHER READING ON PART VII, THE INDUSTRIAL ORGANIZATION

I. Risk, Insurance, and Speculation

Huebner, S. S. *Life Insurance*. New York, 1915
— *Property Insurance*. New York, 1920
Gephart, W. F. *Principles of Insurance (Fire and Life)*, 2 volumes. New York, 1917 1920
Winter, W. D. *Marine Insurance*. New York, 1919
Huebner, S. S. *Marine Insurance*. New York, 1922
Brace, H. H. *Value of Organized Speculation* Boston, 1913
Meeker, J. E. *The Work of the Stock Exchange* New York, 1930
Huebner, S. S. *The Stock Market*. New York, 1934
Boyle, J. E. *Speculation and the Chicago Board of Trade*. New York, 1920
Nourse, E. G. *The Chicago Produce Exchange*. Boston, 1918
Hubbard, W H. *Cotton and the Cotton Market*. New York, 1923
Emery, H. C. *Speculation in the Stock and Produce Exchanges of the United States*. New York, 1896
Baer, J. B., and Woodruff, G. P. *Commodity Exchanges*. New York, 1929
Dice, E A. *The Stock Market*. Chicago, 1927
Knight, F. H. *Risk, Uncertainty, and Profit*. New York, 1933
Bernheim, A. L., and others. *The Security Markets*. Twentieth Century Fund, Inc. New York, 1935

II. Business Cycles and Their Control

PIGOU, A. C. *Industrial Fluctuations.* London, 1927

MILLS, F. C. *Economic Tendencies in the United States.* National Bureau of Economic Research, Inc. New York, 1932

WARREN, G. F., and PEARSON, F. A. *Gold and Prices.* New York, 1935

PERSONS, W. M. *Forecasting Business Cycles.* New York, 1931

TIMOSHENKO, V. P. *The Rôle of Agricultural Fluctuations in the Business Cycle.* (Michigan Business Studies, Vol. II, No. 9), University of Michigan, Ann Arbor, 1930

HANEY, L. H. *Business Forecasting.* Boston, 1931

KUZNETS, S. S. *Cyclical Fluctuations.* New York, 1926

—— *Secular Movements in Production and Prices.* Boston, 1930

MITCHELL, W. C. *Business Cycles. The Problem and Its Setting.* National Bureau of Economic Research, Inc. New York, 1930

HEXTER, M. B. *Social Consequences of Business Cycles.* Boston, 1925

HASTINGS, H. B. *Costs and Profits.* Boston, 1923

ADAMS, A. B. *Economics of Business Cycles.* New York, 1925

—— *Analysis of Business Cycles.* New York, 1936

MCCRACKEN, H. L. *Value Theory and Business Cycles.* New York, 1933

WAGEMANN, E. *Economic Rhythm.* New York, 1930

FISHER, I. *Booms and Depressions.* New York, 1932

MILLS, F. C. *The Behavior of Prices.* National Bureau of Economic Research, Inc. New York, 1927

HANSEN, A. H. *Business-Cycle Theory.* Boston, 1927

—— *Economic Stabilization in an Unbalanced World.* New York, 1932

—— *Full Recovery or Stagnation?* New York, 1938

HABERLER, G. VON. *Prosperity and Depression.* League of Nations, 1937

CLARK, J. M. *Strategic Factors in Business Cycles.* New York, 1934

—— *Economics of Planning Public Works.* New York, 1935

MEADE, J. E. *Economic Analysis and Policy.* Oxford, 1936

ROBINSON, JOAN. *Introduction to the Theory of Employment.* New York, 1937

KING, W. I. *The Causes of Economic Fluctuations.* New York, 1938

SLICHTER, S. H. *Towards Stability.* New York, 1934

ROBBINS, L. *The Great Depression.* London, 1934

DOUGLAS, P. H. *Controlling Depressions.* New York, 1935

III. Problems of the Corporation

(See also suggested readings on Corporations at end of Part I)

JONES, E. *The Trust Problem in the United States.* New York, 1929

SEAGER, H. R., and GULICK, C. A. *Trusts and Corporation Problems.* New York, 1929

LAIDLER, H. W. *Concentration of Control in American Industry.* New York, 1931

JENKS, J. W., and CLARK, W. E. *The Trust Problem.* New York, 1929

DURAND, E. D. *The Trust Problem.* Cambridge, 1915

WATKINS, M. W. *Industrial Combination and Public Policy.* Boston, 1927

Fetter, F. A. *The Masquerade of Monopoly.* New York, 1931
Haney, L. H. *Business Organization and Combination.* New York, 1913
Berle, A. A., Jr., and Means, G. C. *The Modern Corporation and Private Property.* New York, 1933
Henderson, G. C. *The Federal Trade Commission.* New Haven, 1925
Meade, E. S. *Trust Finance.* New York, 1903
Ripley, W. Z. (Editor). *Trusts, Pools, and Corporations.* Boston, 1905
Dewing, A. S. *The Financial Policy of Corporations.* Third edition. New York, 1934
Stevens, W. S. (Editor). *Industrial Combination and Trusts.* New York, 1913
National Industrial Conference Board, Inc. *Mergers and the Law.* New York, 1929
Filene, L., and Stoddard, W. L. *Unfair Trade Practices: How to Remove Them.* New York, 1931
Michaels, R. K. *Cartels, Combines and Trusts in Post-War Germany.* New York, 1928
Liefman, R. *International Cartels, Combines and Trusts.* London, 1927
— — *Cartels, Concerns and Trusts.* New York, 1932
Plummer, A. *International Combines in Modern Industry.* London, 1934
Meakin, W. *The New Industrial Revolution.* New York, 1928
Brady, R. A. *The Rationalization Movement in German Industry.* University of California Press, 1933
United States Commissioner of Corporations — Special Report on the Beef Industry, Steel Industry, International Harvester Company, The Standard Oil Company, etc.
United States Federal Trade Commission — Special Reports on Various Industries. Utility Corporations: Summary Report of the Federal Trade Commission to the Senate. 70th Congress, 1st Session, Document 92, Part 72-A.
Flynn, J. T. *Security Speculation.* New York, 1934
Corey, L. *The Decline of American Capitalism.* New York, 1934
Bonbright, J. C., and Means, G. C. *The Holding Company.* New York, 1932

IV. Public Utilities

A. Railway Transportation

Daggett, S. R. *Principles of Inland Transportation.* New York, 1934
Jones, E. *Principles of Railway Transportation.* New York, 1924
Acworth, Sir W. M. *The Elements of Railway Economics.* Revised edition. Oxford, 1924
Vanderblue, H. B., and Burgess, K. F. *Railroads: Rates-Service-Management.* New York, 1923
MacGill, C. E., and others. *History of Transportation in the United States before 1860.* Washington, 1917
Cunningham, A. J. *American Railroads: Governmental Control and Reconstruction Policies.* Chicago, 1922

Dixon, F. H. *Railroads and the Government.* New York, 1922

Lochlin, D. P. *Railway Regulation since 1920.* Revised edition. Chicago, 1938

Splawn, W. M. W. *Government Ownership and Operation of Railroads.* New York, 1928

Haney, L. H. *The Business of Railway Transportation* New York, 1924

Moulton, H. G., and associates. *The American Transportation Problem.* Washington, 1933

Sharfman, I L. *The American Railroad Problem.* New York, 1921

—— *The Interstate Commerce Commission.* New York, Parts I and II, 1931; Part III A, 1935; Part III B, 1936; Part IV and Conclusion, 1938

Johnson, E R *Government Regulation of Transportation.* New York, 1938

Middletown, P. H. *Railways of Thirty Nations.* New York, 1937

—— *Government versus Private Ownership.* New York, 1937

B. Ocean Transportation

National Industrial Conference Board, Inc. *The American Merchant Marine Problem.* New York, 1929

Berglund, A. *Ocean Transportation.* New York, 1931

Calvin, H. C, and Stuart, E. G. *The Merchant Shipping Industry.* New York, 1925

Johnson, E. R., and Huebner, G. G. *Principles of Ocean Transportation.* New York, 1919

Marvin, W. L. *The American Merchant Marine, Its History and Romance from 1620 to 1902.* New York, 1902

Bates, W. W. *The American Marine.* New York, 1902

Zeis, P. M. *American Shipping Policy* Princeton, 1938

C. Other Public Utilities

Jones, E., and Bigham, T. C. *Principles of Public Utilities.* New York, 1931

Mosher, W. E., and Crawford, F. G. *Public Utility Regulation.* New York, 1933

Wilson, G. L., Herring, J. N., and Eutsler, R. B. *Public Utility Regulation.* New York, 1938

Barnes, I. R. *Cases on Public Utility Regulation.* New York, 1938

Bauer, J, and Gold, N. *Public Utility Valuation for Purposes of Rate Control.* New York, 1934

Graham, W. J. *Public Utility Valuation: Reproduction Cost as a Basis for Depreciation and Rate-Base Determination* Chicago, 1934

Dorau, H. B. *Materials for the Study of Public Utility Economics.* Chicago, 1930

Rauschenbush, S. *The Power Fight.* New York, 1932

V. The Farm Problem

Black, J. D. *Introduction to Production Economics.* New York, 1926

—— *Agricultural Reform in the United States.* New York, 1929

Black, J. D., and Black, A. G. *Production Organization.* New York, 1929

DAVIS, H. A. *Wheat and the AAA.* Brookings Institution, Washington, 1935
BLACK, J. D. *The Dairy Industry and the AAA.* Brookings Institution, Washington, 1935
ROWE, H. B. *Tobacco under the AAA.* Brookings Institution, Washington, 1935
RICHARDS, H. I. *Cotton under the Agricultural Adjustment Act: Developments up to July, 1934.* Brookings Institution, Washington, 1934
NOURSE, E. G., DAVIS, J. S., and BLACK, J. D. *Three Years of Agricultural Adjustment Administration.* Brookings Institution, Washington, 1937
National Industrial Conference Board. *American Agricultural Conditions and Remedies.* New York, 1936
ROBERTSON, C. J. *World Sugar Production and Consumption.* London, 1934
DOWELL, A. A., and JENNESS, O. B. *The American Farmer and the Export Market.* University of Minnesota Press, 1934
TIMOSHENKO, V. P. *World Agriculture and the Depression.* Michigan Business Studies, Vol. V, No. 5, University of Michigan, Ann Arbor, 1933
COPELAND, M. T. *A Raw Commodity Revolution.* Boston, 1938

VI. CONTROL OF COMPETITIVE INDUSTRY

LYON, L. S., and associates. *The National Recovery Administration.* Brookings Institution, Washington, 1935
HACKER, L. M. *A Short History of the New Deal.* New York, 1935
ADAMS, A. B. *Our Economic Revolution.* University of Oklahoma Press, Norman, 1933
BROWN, D. V., and associates. *The Economics of the Recovery Program.* New York, 1934
BERLE, A. A., JR., and others. *America's Recovery Program.* New York, 1934
TAGGART, H. F. *Minimum Prices under the NRA.* Ann Arbor, 1936
ROOS, C. F. *NRA Planning.* Bloomingdale, Indiana, 1937
Brookings Institution. *The Recovery Problem in the United States.* Washington, 1936
FAIRCHILD, FURNISS, BUCK, and WHELDEN. *A Description of the New Deal.* Revised edition. New York, 1935
LUCAS, A. F. *Industrial Reconstruction and the Control of Competition.* New York, 1937
MACLAURIN, W. R. *Economic Planning in Australia, 1929-1936.* London, 1937

VII. GENERAL PROBLEMS OF THE MODERN INDUSTRIAL ORGANIZATION

SCHMOLLER, G. *The Mercantile System.* New York, 1896
ROBERTSON, D. H. *The Control of Industry.* New York, 1923
HOBSON, J. A. *The Evolution of Modern Capitalism.* London, 1917
HADLEY, A. T. *The Conflict between Liberty and Equality.* Boston, 1925
VEBLEN, T. *The Theory of Business Enterprise.* New York, 1904
—— (edited by Ardzrooni, L.). *Essays in Our Changing Order.* New York, 1934
MILL, J. S. *On Liberty.* London, 1859

BENN, E. J. P. *Return to Laissez Faire*. New York, 1929

KEYNES, J. M. *End of Laissez Faire*. London, 1926

CHASE, S. *The Economy of Abundance*. New York, 1934

HAMILTON, W., and others. *Prices and Price Policies*. New York, 1938

LYON, L. S., and ABRAMSON, V. *The Economics of Open Price Systems*. Brookings Institution, Washington, 1936

BURNS, A. R. *The Decline of Competition*. New York, 1936

NOURSE, E. G., and associates. *America's Capacity to Produce*. Brookings Institution, Washington, 1934

LEVEN, M., MOULTON, H. G., and WARBURTON, C. *America's Capacity to Consume*. Brookings Institution, Washington, 1934

MOULTON, H. G. *The Formation of Capital*. Brookings Institution, Washington, 1935

PERSONS, W. M. *Government Experimentation in Business*. New York, 1934

DIMOCK, M. E. *Government-Operated Enterprises in the Panama Canal Zone*. University of Chicago Press, 1934

—— *British Public Utilities and National Development*. London, 1933

Power Authority of the State of New York. Report on Cost of Distribution of Electricity. Albany, 1934

THOMPSON, C. D. *Public Ownership*. New York, 1925

MEYER, H. R. *Municipal Ownership in Great Britain*. New York, 1906

HIBBARD, B. H. *A History of the Public Land Policies* New York, 1924

CRECRAFT, E. W. *Government and Business*. New York, 1925

REQUA, M L. *Relations of the Government to Industry* New York, 1925

SPLAWN, W. M. W. *Government Ownership and Operation of Railroads*. New York, 1928

KNOOP, D. *Principles and Methods of Public Trading*. London, 1912

O'BRIEN, T. H. *British Experiments in Public Ownership and Control*. New York, 1938

SORRELL, L. C. *Government Ownership and Operation of Railways for the United States*. New York, 1937

PART VIII

PROBLEMS OF LABOR

XLVIII

POPULATION PROBLEMS

The problem of population has both a quantitative and a qualitative aspect. On the one hand, there is the question of the number of people in relation to the supporting power of their physical and institutional environment; on the other, the question of the capabilities and social worth of the individuals composing the group. Interest in the qualitative side of the problem has been of fairly recent growth. It has given rise to theories of the superiority of certain racial stocks and to a group of policies — such as the eugenics program — designed to improve the physical and moral qualities of the population by controlling the breeding of future generations. Students of population problems were at first concerned merely with questions of numbers and rates of increase, and this quantitative aspect continues even today to receive major emphasis.

Malthusian theory of population: Statement. The question of population increase and the social problems resulting therefrom were first handled in a systematic manner by Thomas Robert Malthus. Malthus' book, *An Essay on the Principle of Population*, published in 1798, made an immediate and lasting impression on all students of population problems, so much so that the term "malthusianism" continues even now to be applied to this field of inquiry.

Examination of a large mass of data regarding actual rates of population increase led Malthus to the following conclusions: (1) population is necessarily limited by the means of subsistence; (2) population invariably increases where the means of subsistence increase, unless prevented by some very powerful and obvious checks; (3) the checks which keep population increase down to the level of the means of subsistence may be classified into three categories: moral restraint, vice, and misery. It was the general infer-

ence of Malthus' teaching that increase of population, pressing continuously upon the supporting power of the soil, must prevent any large improvement in average material welfare. In particular any attempt to improve the condition of the mass of the people through social reform must be defeated by the inevitable increase of numbers whenever conditions of life are eased. It was the second of Malthus' three propositions – namely, that the tendency of population to expand is "invariable" — which gave chief support to these gloomy views of the problem.

The Malthusian argument is based on a comparison of the rate of population increase with that of the increase of food. His historical studies had convinced Malthus that population when unchecked increases in geometrical ratio, doubling every twenty-five years; while food increases in arithmetical ratio. For example population would tend to increase in two centuries as the numbers 1, 2, 4, 8, 16, 32, 64, 128; but food supplies would increase as the numbers 1, 2, 3, 4, 5, 6, 7, 8. Hence there must be increasing intensity of pressure of population on the means of subsistence, resulting in misery and general degradation But since it is physically impossible for population actually to outstrip the food supply, forces must be in operation to restrain the growth of numbers. These "checks" are of two different types: the preventive check, "moral restraint," which reduces the birth rate; and the positive checks, "vice and misery," which increase the death rate. Malthus presents these as alternative choices of mankind; unless the preventive check is voluntarily applied by the human race, expansion of population must inevitably bring the positive check into operation. Starvation, disease, infant mortality, and war are the ultimate restraints on population growth, and these manifestations of "vice and misery" cannot be evaded unless something is done to reduce the fecundity of the people. Malthus himself favored an increased use of "moral restraint" but thought it unlikely that people in general would fall in with his views; hence the gloomy nature of his outlook on the future.

Appraisal of Malthus' theory. Malthus' reasoning is open to criticism at several points. His comparative ratios of the increase of population and of food are hypothetical and lacking in probative

value. Even if his definition of the ratios is correct, it should be observed that a geometric ratio may result in so slow a rate of increase and an arithmetic ratio in so rapid a rate that the one will not overtake and outstrip the other within a period of time brief enough to be of human concern.

Moreover Malthus was in error in assuming that population would necessarily expand at a rate determined by the extreme fecundity of the human race. He mistook a biological possibility for a human tendency. The potentialities of increase are undoubtedly present, but many influences of a non-biological character intervene to slow down the rate of increase.

Furthermore the theory leaves out of account the possibility of an advance in the arts of production. It assumes that each succeeding generation of people must win their subsistence from the soil with the aid of no arts unknown to their predecessors. This we know is not the case; the rapid technological progress of the past century and a half has started each generation of men on a new and higher economic plane. It is conceivable that this increase of productive power may proceed so rapidly as to postpone indefinitely the pressure of numbers on the supporting power of the soil. Safe prediction on this point must be based, not on rules of logic and a comparison of hypothetical ratios, but on accurate knowledge of the facts of economic progress and of the actual increase in population.

But when all this has been said, the Malthusian doctrine still contains a truth of great importance. In view of the fecundity of the human race, it cannot be denied that if no restraint is placed upon the birth rate and no great advance of technology occurs, the number of human beings will presently outstrip the consumable income of society, causing degradation of the standard of life. This results from the operation of the law of diminishing returns in agriculture and in all forms of industry devoted to the improvement of raw materials taken from nature. On a given stage of the arts there is a certain density of population which will be the most efficient from the point of view of return per head. If population has not attained this number, every additional man will increase the productive effectiveness of the group and raise the average

income. On the other hand, if it passes beyond this number, each new producer will add a smaller amount to the social product, and the average income will fall. Considering that human increase may at any time become an almost automatic biological process with vast powers of expansion, and that progress in the arts of production is continuously retarded by the limitations of human intelligence, the race of men is at all times confronted with the possibility of a redundancy of numbers. Stating Malthus' doctrine in more cautious terms, it is true that population has the power to increase up to the limit of the supporting power of the environment on a given stage of the arts. This tendency must inevitably cause suffering, misery, and vice (the positive checks of Malthus) unless either it is held in check by control of births (the preventive check) or its effects are postponed by timely advances in science and the arts of production.

Overpopulation still a menace. The fortunate experience of many countries during the nineteenth century, when population expanded rapidly and the standard of living at the same time rose, does not really dispose of the Malthusian problem. On the one hand, it should be noted that in many parts of the world — in India and China, for example — the forebodings of Malthus have been fulfilled. Here population presses insistently upon the supporting power of the soil, and starvation, misery, and a high death rate are evidences of the inexorable operation of the positive check. Moreover even among European peoples the developments of the decades just past would indicate that the era of increasing population coupled with improved living conditions is at an end. At this moment the real problem is blurred in many European countries by dictatorship, which, on the one hand, fosters an increasing population for military reasons and, on the other, conceals the actual effect of present numbers upon average economic well-being by exalting patriotism and race pride as a recompense to poverty. But actually standards of living have begun to fall, and the danger of overpopulation may again become a problem of grave practical importance.

Many modern students of the problem are convinced that there will be no return of the conditions which made possible the un-

precedented expansion of European peoples during the nineteenth century. Some improvement may be made in agricultural techniques and in transportation. But the amount of arable land drawn into use is approaching its maximum. Efforts to increase the food supply must be concentrated on a definitely restricted land area, taking the form of greater intensity of cultivation, with its consequence of lower average output per worker. This condition will become increasingly apparent in the more highly industrialized nations as the increase of population in undeveloped countries absorbs a larger proportion of the foodstuffs and raw materials which these undeveloped countries produce. Already Europe fails to produce enough food to feed her people. The New World from which Europe draws the remainder of her food supply is increasing rapidly in population, and it also is becoming industrialized so that it is less desirous of accepting manufactures from Europe in exchange for food. It seems inevitable that countries which are not now self-supporting as regards food and raw materials must experience either a decline in the standard of living or a restraint of population to a point comparable with their natural productivity.

Present-day examples of overpopulation are to be found among the peoples with the highest birth rates. Russia, British India, China, and French Canada are cases in point. Here the average birth rates are about 50 per thousand inhabitants, while in some localities it mounts as high as 55 or 60. Such birth rates mean that the population is reproducing fast enough to double every twenty-five years, or to expand in a century to sixteen times its present size But among these same peoples the high birth rates are coupled with abnormal death rates. In China and India, for example, the death rate is probably 50 per thousand; that is, equal to the birth rate among these people. In other words, the population is actually almost stationary; the average span of life is only about twenty years; deaths cancel births with an enormous wastage of human life, and the struggle for existence reduces even the survivors to a very low level of life. These are infallible signs of *de facto* overpopulation.

Among the more progressive peoples of the world, the very improvements in culture on which they pride themselves as proofs

of their high standards of life may lead eventually to overpopulation. This is particularly true of advances in the medical arts, which have caused a sensational lowering of the death rate among the more advanced nations. During the third of a century preceding the World War these developments had reduced the rate of infant mortality in progressive countries by an amount varying between a quarter and two-fifths. Such progress in public health and medical science must accelerate the growth of population unless counteracted by a simultaneous decline of the birth rate. That birth rates have declined is shown at a later point in this chapter, but in many countries the reduction of the death rate has been more rapid than the decline of births.

Death rates. The lowest death rates so far reached in the countries where records are kept are those of Holland, Great Britain, Australia, the Scandinavian countries, and some of the American states. The highest rates are those of Russia and Rumania in Europe and Chili in America. The following table records the death rate in certain countries in 1920.

ANNUAL CRUDE DEATH RATES IN 1920

	Deaths per 1,000 population
New Zealand	8 7
Australia	9 9
United States	11 8
Netherlands	12 8
Wales	12 8
Denmark	12 9
Norway	13 2
England	13 7
Sweden	13 8
Switzerland	14 3
Belgium	15 2
Scotland	15 4
Germany	16 2
Ireland	17 6
France	18 4
Italy	19 2
Japan	21 9
Austria	21 9
Rumania	23 4
Bulgaria	26.4
Chili	31 1

A high death rate frequently goes with a high birth rate, because both are characteristics of impoverished conditions of life. The same conditions that favor excessive fecundity make inevitable a high rate of mortality.

The possible decline in the death rate is of course limited. In China, with an average birth and death rate of probably 50 per thousand, the average span of life is about twenty years. A society with a death rate of 35 would have an average span of life of under thirty years. If the death rate be reduced to 20, it would mean an average life of fifty years. If it be reduced to 10, it would mean that every child born would live to be one hundred, or if some die younger, others would need to live longer. It is evident that a death rate of approximately 9, as in New Zealand, or of approximately 10, as in the Australian Commonwealth, cannot long be maintained. Such rates, in the absence of a large immigration of persons in the safer years of life, are destined to rise. It is equally evident that a death rate of 25 or 30 is excessive and that such rates would fall rapidly with the extension of education and the introduction of modern methods of sanitation.

The declining birth rate. The rapid decrease of births in many countries has caused much more discussion than the even larger decline of the death rate. About seventy years ago the population of France began to show a sharp decrease in fertility. In the next decade the same phenomenon was observed in England. At this time wide publicity was being given to an attempt on the part of the British government to punish a group of people for spreading knowledge of birth control The most important effect was the immediate popularization of this proscribed knowledge, with the result that the birth rate in England fell by as much as one-third in the next thirty-five years. From France and England the decline in birth rates spread to all countries of Europe, to Australasia, and to the United States.

The table on page 450 gives a comparison of the birth rates of certain countries at different periods of time.

It should be noted that in spite of the decline in the birth rate since the seventies, the populations of practically all countries experiencing this change have rapidly increased. This is owing to

the larger population to which the rate, though reduced, is applied and to the fact that the death rate has declined even more rapidly than the birth rate. The situation in France, where there is now a stationary if not an actually decreasing population, is due not solely to the very low birth rate but quite as much to the fact that the death rate has not been lowered there as rapidly as in other countries.

BIRTH RATE PER 1,000 POPULATION

	1871–1880	*1901–1910*	*1920*
Australia	36.1	26 5	24.9
Austria	39.0	34.7	31.4
Belgium	32.3	26 1	23 7
Denmark	31.4	28 6	25.9
England	35.5	28 2	23 1
France	25 4	20.6	18.7
Germany	39.1	32.9	29.8
Italy	36.9	32 7	31.3
New Zealand	40.5	26 8	23 3
Norway	31.0	27.4	25.9
Sweden	30.5	25.8	23.8
Switzerland	30.7	26.9	22.7

Population growth in the United States. In the early decades of American history the rapidity with which the population increased through births made this country a favorite example of the Malthusian principle that people will multiply at a great rate when food is plentiful. Later the vast influx of immigrants coöperated with the high birth rate to promote population expansion. However recent studies of successive census returns show that during most of this period the rate of increase of the population through births was steadily declining, and now that immigration is no longer a factor of importance the absolute increase is declining also. At the present time the increase is about 800,000 a year as against 1,800,000 ten years ago. If this trend continues, the population will become stationary at a comparatively early date, not later than 1960.

The decline of the ratio of births to the number of women of child-bearing age has been substantial since the middle of the last century, and particularly marked since 1920. At the close of the nineteenth century the white women of this country were bearing between 350 to 400 children per thousand women during their

fertile period; an average of nearly 4 births to each adult woman. By 1930 the average per adult white woman had fallen to 2.2. If the present trend continues until 1955 the rate in that year will have fallen below 2. Apart from its effect on numbers, this trend of the birth rate is changing materially the composition of the population. The average age of the people is increasing. In 1870, half of the people were under twenty years of age, 45 per cent were between twenty and sixty years of age, 5 per cent above sixty years; in 1930 the proportion in the first age group had fallen to less than 40 per cent, that in the second had risen to over 52 per cent, and that in the third to 9 per cent. By 1950, if this trend continues, only 30 per cent of the people will be under twenty years of age, 57 per cent will be between twenty and sixty years of age, and 13 per cent above sixty years. These changes may have important consequence in national life. Political activities and the attitude on social problems will be affected by the shifting of preponderant voting strength into the higher age groups. The economic situation will be influenced by the increase in the proportion of people in the productive years of life — twenty to sixty years of age — as compared with the dependent groups, the very old and the young.

It may be observed too that, despite the urban growth of the country, the population continues to be recruited in large part from the farms. At the present expectation of life in the United States, 360 children under five years of age per 1,000 women of child-bearing age are required to maintain a stationary population. In 1930 all the cities over 100,000 population fell 20 per cent short of their quota, and the smaller cities showed a 5 per cent deficit, while the non-urban farm population contributed a surplus of 31 per cent and the farming population a surplus of 51 per cent.

The differential birth rate. Special problems are presented by the fact that the rate of increase varies on the different social levels of the population and among different racial stocks. Throughout the civilized world vital statistics show a high correlation between the birth rate and the conditions which indicate a lower social status, such as the amount of child employment, pauperism, overcrowding, and other indications of poverty and lack of culture. The higher the social status the lower is the fertility of the people,

and this difference in rate of increase is not counterbalanced by difference in death rates. The United States shares in this experience, and it is everywhere a recent phenomenon.

A similar correlation is found between the birth rate and the degree of education or the intellectual capacity of the population groups. Studies in the United States show that the birth rate is declining among college-bred men and women much more rapidly than among people with less education. The average graduate of Vassar is the mother of one-half of a daughter and the average graduate of Harvard the father of three-fourths of a son. Similarly the graduates of Yale and the other universities fail to replace their numbers through their offspring. Statistical studies of birth rates among different professional groups show like results. Among American men of science, for example, the average family has two children, although the average for the country as a whole is about four and a half.

In the United States the differential birth rate has further significance in that the immigrant stock is more prolific than the native American. Studies of the Immigration Commission showed an average family of 2.7 children for native white women of native parentage married from ten to twenty years, while for white women of foreign parentage in the same age group the average was 4.4 children per family. Of women of native parentage in this group, 13 per cent were childless as against 5.7 per cent of foreign-born women. Only 10 per cent of the native-born had five or more children, while 33 per cent of the foreign-born had families of this size. A later study of the 1930 census shows that cities whose stock is chiefly American-born lacked by 38 per cent enough children to maintain their present populations. Judging from all the evidence available, the rate of increase in the alien racial groups of the population is twice as rapid as in the native stock.

Birth control. There can be no doubt that the general decline of the birth rate has been due in large part to intentional control of births by married people and that the differential rate indicates a wider adoption of these practices by certain social classes and racial stocks than by others. The spread of birth control throughout the nations of western Europe and America during the past half-

century is one of the striking social phenomena of the time. At its beginning this movement was almost universally condemned by governmental authority, and even today the dissemination of information regarding contraceptive devices is illegal in most other countries, as it is in the United States. The weight of organized religion has been thrown consistently against the practice without preventing its spread, although this influence has perhaps been a factor in accentuating the differential birth rate among various social classes. In general the people of these countries have applied the "preventive check" to population on an increasing scale as the years have passed.

The motives for limiting the size of the family are many. In quality they run the gamut from altruistic to selfish. They may be summarized as follows: (1) the economic burden of children, especially in urban environments, because of the high cost of living, poverty, employment of women, child labor laws, the lengthened period of compulsory school attendance, etc., (2) increasing regard for child life, its nurture and education, and a keener sense of parental responsibility, (3) desire for social position, attributed to a selfish love of luxury, desire for comfort and economic security, vulgar ambition, prudence, etc., (4) changing position of women as regards increased liberty and economic independence, dislike of child-rearing, avoidance of the pain of childbirth, (5) the increasing demand for long preparation for the professions and other occupations requiring considerable training, which leads to late marriages, and so on, most of the factors being resolvable into a rising standard of living.

It is not our purpose to enter the controversy regarding the practice of birth control. Until recently the arguments against contraceptive practices were almost wholly religious or moral. In our own day the opposition has been strengthened by the outburst of intense nationalism in many countries, which has led to a condemnation of birth control as a betrayal of the nation's future welfare in case of war. These religious, moral, and patriotic aspects of the question will not be further considered here, not because they are unimportant, but because their non-economic character places them beyond the scope of our study. It may be said in summary of

the arguments on the other side that they derive in general from the principle that a lessened pressure of population on the resources of the world means economic progress. But such generalizations have actually little to do with the matter. Personal motives and desires supply the determining reasons for birth control by those who practice it.

The question of quality of population. As has been said, this aspect of the problem of population has attracted serious attention only within comparatively recent times. The differential birth rate more than anything else has aroused interest in the question. It is obvious that the groups with the highest birth rates must tend to crowd out the less fertile groups. Moreover if the expanding groups possess inferior characteristics of a sort transmissible by inheritance, the result must be a gradual degradation of the quality of the population.

The United States is recruiting its numbers chiefly from the lower income groups, from the social classes with relatively little education and culture, from racial stocks which are alien to those which first peopled the country and established its institutions. Whether or no this spells degradation depends on whether the children of the poor and the uneducated are less well endowed in natural abilities than those of the well-to-do and cultivated classes and whether one racial stock can be shown to be innately superior to another. On these questions there is too little scientific evidence to permit of dogmatic answers. With regard to racial differences particularly, much of the recent literature exalting the social virtues of one race — *e.g.*, the "Nordic" — and decrying the inferiorities of the others is too emotional and unscientific to be of value. The trend of scholarly opinion has been to minimize the differences in innate racial ability popularly assumed to exist. However when we consider the possible cultural effects of the differential birth rate, rather than its effect on the innate quality of the population, there is room for serious misgiving. There are great economic and cultural differences between the later immigrant stocks and the older native population. To submerge the native stocks through excessive increase of these alien groups will certainly alter the character of institutional and cultural life, perhaps not for the better.

Eugenics. The work of Sir Francis Galton, who coined the word "eugenics," gave rise to an organized movement to improve the physical quality of the population by selective breeding. The policies proposed have both a positive and a negative side; on the one hand, the attempt is made to increase the birth rate of superior family strains; on the other, to reduce that of inferior strains and in extreme cases to prevent these strains from having offspring.

The positive side of the eugenics program involves assumptions which are not always tenable. In the first place, it often assumes that the classes with the lowest birth rates are biologically best. Secondly it implies that exceptional capacity is a family trait rather than an individual variation, a matter which has not yet received scientific verification. Thirdly and more important, it fails to distinguish the cultural from the biological aspects of the problem. Race improvement and race deterioration are biological concepts; civilization — the progress of culture — is a social concept. Social and cultural improvement are not matters of biology and not necessarily contingent on race improvement; conversely biological improvement gives no necessary assurance of social advance.

Not only is there the problem of discovering just who are the "fit", there is also the question as to what we should breed for. Even if we know how to produce children with certain characteristics, we are not yet agreed as to what characteristics we want. As Dean Inge says: "The two ideals, that of the perfect man and that of the perfectly organized state, would lead to very different principles of selection. Do we want a nation of beautiful and moderately efficient Greek gods, or do we want human mastiffs for policemen, human greyhounds for postmen, and so on?" Many of us would answer that we do not want either the one or the other; we would much rather have the present varieties of human beings.

Negative eugenics, on the other hand, presents a more definite and practical program. Its object is to stop the increase of certain extreme forms of biological unfitness, especially where the defect is due to heredity. Measures such as segregation, sterilization, prohibition of marriage, and the like have been suggested and in certain

respects put into practice. Most of our states now have laws denying the right of marriage to certain classes of defectives, such as the feeble-minded, insane, and epileptic; most have hospitals and institutions where such persons may be confined; while some have laws providing for the sterilization of such individuals. These laws are not adequately enforced, nor will they be until states come to make an examination of the total population and a registration of all defective members. The measures applied by these laws are sometimes denounced on sentimental and emotional grounds. From a rational viewpoint however there is little to be said against them, provided care be taken to apply them only to the biological — not the cultural — defectives and then only in extreme cases.

Immigration to the United States: Statistics. In the United States problems of population have been greatly complicated by the continuous influx of new inhabitants from other countries. Official statistics of immigration have been kept since 1820. It has been estimated that from 1776 to 1820 the total number of arrivals was 250,000. Between 1820 and 1934, over 37,000,000 immigrants have been admitted. The following table lists the influx by decades.

TOTAL IMMIGRATION INTO THE UNITED STATES BY DECADES

1820-29	128,502
1830-39	538,381
1840-49	1,427,337
1850-59	2,814,554
1860-69	2,081,261
1870-79	2,742,137
1880-89	5,248,567
1890-99	3,694,294
1900-09	8,202,388
1910-19	6,347,380
1920-29	4,295,510
1930-37	548,482
Total	38,068,794

Beginning with 1908, the Bureau of Immigration has kept records of departures as well as of arrivals. It is thus possible to secure figures of net immigration for this period.

Character of immigration. The history of immigration to the United States may be divided into two major periods: (1) the period of the so-called old immigration, from the beginning of our inde-

pendent national existence to 1882, during which time our immigrants came chiefly from the British Isles, the Scandinavian countries, and Germany, and (2) the period of the new immigration, from 1882 to the present, in which most of the immigrants have come from Italy, former Austria-Hungary, Russia, and other countries in the south and east of Europe. This shifting of the main sources of immigration from northern and western Europe to southern and eastern Europe is the most outstanding fact in the history of the movement to this country. It has meant, in general terms, the coming of peoples who are more alien to the native stock as to race, language, history, religion, customs, and traditions.

NET IMMIGRATION INTO THE UNITED STATES, 1908–1937

	Immigrants admitted	*Immigrants departed*	*Net immigration*
1908	782,870	395,073	387,797
1909	751,786	225,802	525,984
1910	1,041,570	202,436	839,134
1911	878,587	295,666	582,921
1912	838,172	333,262	504,910
1913	1,197,892	308,190	889,702
1914	1,218,480	303,338	915,142
1915	326,700	204,074	122,626
1916	298,826	129,765	169,061
1917	295,403	66,277	229,126
1918	110,618	94,585	16,033
1919	141,132	123,522	17,610
1920	430,001	288,315	141,686
1921	805,228	247,718	557,510
1922	309,556	198,712	110,844
1923	522,919	81,450	441,469
1924	706,896	76,789	630 107
1925	294,314	92,728	201,586
1926	304,488	76,992	227,496
1927	335,175	73,366	261,809
1928	307,255	77,457	229,798
1929	279,678	69,203	210,475
1930	241,700	50,661	191,039
1931	97,139	61,882	35,257
1932	35,576	103,295	−67,719
1933	23,068	80,081	−57,013
1934	29,470	39,771	−10,301
1935	34,956	38,834	−3,878
1936	36,329	35,817	+512
1937	50,244	26,736	+23,508
Total	12,726,028	4,401,797	8,323,207

A line drawn across the continent of Europe from northeast to southwest, separating the Scandinavian Peninsula, the British Isles, Germany, and France from Russia, former Austria-Hungary, Italy, and Turkey, separates countries not only of distinct races but also of distinct civilizations. It separates countries of representative institutions and popular government from those accustomed to absolute monarchies; it separates lands where education is universal from lands where illiteracy predominates; it separates manufacturing countries, progressive agriculture, and skilled labor from primitive hand industries, backward agriculture, and unskilled labor; it separates an educated, thrifty peasantry from a peasantry scarcely a single generation removed from serfdom; it separates Teutonic races from Latin, Slav, Semitic, and Mongolian races. When the sources of immigration are shifted from the Western countries so nearly allied to America to Eastern countries so remote in the main attributes of Western civilization, the change is one that should challenge the attention of every citizen.[1]

Causes of the changed character of immigration. It is impossible to go deeply into the reasons for this changed character of immigration, but some of the main causes are as follows: The great industrial expansion of the United States has taken place since the eighties, and a study of the facts of the case will show that it has gone hand in hand with the increase of immigration from southern and eastern Europe. America used to be chiefly an agricultural nation; it is now a great manufacturing and commercial nation. While there was free land and while the element of skill was a large factor in industry, conditions were favorable for the immigration of northern and western Europeans. When the free land had practically disappeared and the invention of machinery had reduced the need for skilled men and greatly increased the demand for unskilled workers, the United States needed and got a new type of immigrant. Various other factors, such as the removal of religious and political causes of emigration from northern and western Europe and the fear of competition with recent immigrants of lower standards of living, help to account for the changed character of immigration, but from an economic standpoint the fundamental reasons seem to

[1] J. R. Commons, *Races and Immigrants in America*, p. 70.

be the ones just mentioned. The general conditions among the skilled workers of northern and western Europe are such that there is little motive from that source of emigrating. Demand during the period of the "new immigration" being chiefly for unskilled workers, America got mostly unskilled immigrants, no matter from what section of Europe they came. Southern and eastern Europe could furnish proportionately more unskilled industrial workers than northern and western Europe, because the general standard of living is lower and there is less industrial development; so most of the immigrants came from that section These statements of course do not embrace the most recent refugee immigration from central Europe.

Restrictive legislation. The changed character of immigration has had more influence than any other single factor in developing in the United States a body of opinion favoring restriction of immigration This has recently found expression in the percentage restriction laws, which constitute a radical departure from the former policy of regulating European immigration. Throughout the period during which there has been legislation on the subject of immigration, three different types of laws have been enacted. These involve the principles of exclusion, selection, and restriction.

Exclusion. Outright refusal to admit the people of certain nations as immigrants to this country has been based in part on racial considerations, in part on economic. The desire to protect the native race from further dilution by dark-skinned peoples and to avoid aggravation of already severe racial problems by admitting non-assimilable groups were motives for this policy But even more influential was the demand of the laboring population that they be protected from competition destructive of their standard of living. The plane of life among workers of the black, brown, and yellow races is so low that the American-born whites cannot allow these workers to set the terms of the wage bargain in the labor market. The principle of exclusion has been applied solely to Asiatics. The Chinese were the first to be excluded, and they have been kept out by a series of acts beginning with that of 1882, which suspended the immigration of Chinese laborers. At Japan's request the United

States entered the so-called "gentlemen's agreement" in 1907, a passport arrangement which provided that Japan would not, except under certain conditions, issue to laborers passports good for continental United States. The immigration act of 1924 terminated this informal arrangement and excluded the Japanese by name. Other races were excluded by the immigration act of 1917, which created the Asiatic barred zone, from which no immigrants may come. This area, defined by degrees of longitude and latitude, includes India, Siam, Indo-China, parts of Siberia, Afghanistan, and Arabia, and the islands of Java, Sumatra, Ceylon, Borneo, New Guinea, Celebes, and various smaller groups. Adding Japan and China this is the full list of countries which have so far been dealt with on the principles of exclusion. It should be noted however that exclusion applied in all these instances only to the laboring classes. Travelers, students, teachers, and officials, as well as certain other exempt classes, are still free to come.

Selection. The principle of selection has been applied to European immigrants since 1882, when the first inclusive immigration law was passed by the federal government. It became obvious to everyone that America could not go on admitting all manner of persons without getting most of the dregs of Europe. Indeed many European nations had been making a definite practice of using the United States as a dumping ground for their criminals, paupers, and other undesirables. In self-defense this country had to regulate immigration, and the principle of selection, which imposes qualitative tests for admission, was adopted. The law of 1882 prohibited the entry of convicts (other than political offenders), idiots, lunatics, and persons likely to become public charges. Since then there has been a series of laws of increasing severity, which have added to the list of ineligibles and have become more and more stringent regarding the deportation of those who should have been excluded but who slipped by. The most recent general immigration act, that of 1917, applies a test of physical, mental, moral, and (in a limited way) educational fitness, and to this extent is a selective law. Its purpose is to exclude from admission all aliens who upon examination at American ports are found to fall below the prescribed physical, mental, or moral standard,

those who are diseased, and (with some exceptions) those who cannot read in the English language or some other language or dialect.

Restriction: Demand for restriction. The selective policy places no necessary limitation upon the number of immigrants admitted, as is shown by the figures of the table on page 456 for the decades in which this policy was in effect. Opinion in the country that something should be done to reduce the total number of newcomers grew into a political demand for rigorous restriction. The opinion was supported by a variety of arguments. The huge influx of laborers with a low standard of living produced in the labor market a competitive struggle similar to that caused at an earlier date by the Asiatic immigration. It was feared that unrestricted immigration would eventually reduce the standard of life among the wage earners to the lowest level of European countries that were permitted to send immigrants. Race feeling was aroused by the great diversity of foreign elements which composed the "new" immigration. The ability of a native stock to assimilate a foreign stock is largely determined by the ratio which the former bears to the latter. It became known that two-thirds of the people in the country were of foreign birth or with foreign parentage in one or both parents. When only one-third of the nation could lay claim to full native parentage, assimilation to the American type became extremely difficult. Patriotic feelings aroused by the World War forced the problem of assimilation to the fore. It came as a shock to the people to learn while the country was at war that five of the fourteen million persons of foreign birth could not read, write, or speak our language; that over half the foreign-born whites of voting age were not citizens; and that some European governments were inducing their nationals in this country to retain their old allegiance.

The demand for restriction arising from these sources of opinion found expression in the immigration act of 1921, which limited the number of aliens of any nationality admissible annually to three per cent of the number of persons of that nationality shown to be resident in this country by the census of 1910. This was a temporary act, expiring by limitation June 30, 1924. It was replaced

by the permanent quota limit act of that year, which now forms the basis of the limitation policy.

The quota system. The immigration act of 1924 is an addition to and not a substitute for the earlier acts. In particular the selective features of the law of 1917 continue in force, as does the policy of exclusion from the Asiatic barred zone. In fact exclusion was made more severe by a clause of the act of 1924 providing that "no alien ineligible to citizenship shall be admitted to the United States." Since our state naturalization laws debar all races but the white and the African black, this provision makes more effective the exclusion of Asiatics.

The following countries are not included in the quota system: Canada and Newfoundland, Mexico, Cuba, Haiti, the Dominican Republic, the Canal Zone, and the independent countries of South America. Aside from the barred zone and these exempted areas, immigration from each country of the world is given a maximum annual figure determined by that nation's quota. Furthermore the act provides that after 1927 the total immigration from all sources covered by the quota system shall not exceed 150,000 a year.

Between 1924 and 1929, while permanent measures for determining the quotas were being worked out, temporary devices were used to fix the limits of each nation's contribution to the immigrant stock. The system put into effect in 1929 and still operative is known as the national-origins basis of immigration. The total quota of 150,000 is apportioned among the countries to which the act applies according to their relative contribution to our population as enumerated in 1920, with the proviso that the minimum allotment to any country shall be 100. The determination of these quotas was a very difficult problem, into whose intricacies we need not enter. In brief the process was as follows: first to ascertain the total number of people in continental United States in 1920 whose origin by birth or ancestry was attributable to the countries included under the act; next to determine what fraction of this number had been contributed by each country; finally to apply this ratio to the total quota of 150,000, the result being each nation's annual allotment. If for example 17 per cent of the people in the United States in 1920 having such national origin as above ex-

plained were German by birth or extraction, then 17 per cent of 150,000, or 25,500, would be Germany's annual quota.

It will be seen that this system not only reduces the total immigration to small dimensions, but places the national ratios on a basis quite different from that which was true of recent experience. The old immigration has been favored relatively as compared with the new. This is true because the survey of national origins which formed the basis of the computations included our entire history. For most of this time, the immigration was drawn chiefly from the British Isles, Scandinavia, and Germany; and these nationals, having been here longer, had made cumulative contributions through the birth rate to the population as of 1920. The result is shown by the following table of quotas established in 1929:

Country	*National-origins quotas*
Austria	1,413
Belgium	1,304
Czechoslovakia	2,874
Denmark	1,181
France	3,086
Germany	25,957
Great Britain and North Ireland	65,721
Irish Free State	17,853
Italy	5,802
Netherlands	3,153
Norway	2,377
Poland	6,524
Russia	2,784
Sweden	3,314
Switzerland	1,707

These allotments give to southern and eastern Europeans only 15 per cent of the total quota and to northern and western Europeans over 80 per cent. Between 1910 and 1914 — the most recent normal period of immigration prior to the quota law — the average net annual influx from southern and eastern Europe was 738,000 as against 126,000 from northern and western Europe.

Non-quota immigration. To avoid misunderstanding it may be well to repeat that immigration from many countries, enumerated above, has not been subjected to this restrictive system. Moreover the act specifically provides that certain types of people are to be

admitted from the areas covered by quota restrictions without being chargeable to the quotas of their respective countries. These non-quota immigrants whose numbers are not restricted include the following classifications of people:

(1) An immigrant who is the unmarried child under twenty-one years of age or the wife of a resident citizen of the United States who files a petition stating that he is able and willing to support the immigrant if necessary to prevent such immigrant from becoming a public charge;

(2) An immigrant previously lawfully admitted to the United States, who is returning from a temporary visit abroad;

(3) An immigrant who was born in the Dominion of Canada or in any one of the other countries of the New World mentioned above and his wife and his unmarried children under eighteen years of age, if accompanying him or following to join him;

(4) An immigrant who seeks to enter the United States solely for the purpose of carrying on the vocation of minister of any religious denomination or professor of an institution of learning and his wife and his unmarried children under eighteen years of age, if accompanying him or following to join him;

(5) An immigrant who is a bona fide student at least fifteen years of age and who seeks to enter the United States solely for the purpose of study at an accredited educational institution.

Visa and preference. Each immigrant must be in possession of an immigration visa issued by a United States consul before he can be admitted to the United States, and the annual and monthly limitation under the various quotas is controlled through limiting the number of quota immigration visas issued in any month or year. In other words the quotas are counted or controlled in American consulates, usually in the country where the applicant resides, rather than on arrival at a United States port, thus obviating the necessity of excluding would-be immigrants after arrival here. This method of administration provides easy means of regulating the flow of immigration and also makes possible a certain amount of selection through the preference shown to some applicants.

Preference in the issuance of quota visas is given (1) to a quota immigrant who is the unmarried child under twenty-one years of

age, the father, the mother, the husband, or the wife, of a citizen of the United States who is twenty-one years of age or over, and (2) to a quota immigrant who is skilled in agriculture and his wife and his dependent children under the age of sixteen years, if accompanying him or following to join him. A preference in the case of persons skilled in agriculture is not applicable to immigrants of any nationality the annual quota for which is less than 300, and in no case shall the combined preferences exceed fifty per cent of the annual quota of any nationality.

Social consequences of the restrictive policy. The present American immigration policy is by no means faultless. The quota system has created certain administrative difficulties which have worked hardship in many individual cases; though these minor defects in the law have been and can still be improved by changes of detail. More serious is the practice of discriminating against certain nationalities as a group, implying that all immigrants from these regions are inferior to all who come from the favored nations. This attitude toward the problem of race is fallacious; superiority is an individual matter. Quantitative restriction does not allow for the selection of the better types within each group.

On the whole however the new attitude toward immigration is sound. Mere increase in numbers offers slight economic advantage to the country. It is much more important to improve the quality of the population than to increase its size. The tendency of all peoples to expand their numbers by natural increase under favorable conditions can be relied upon to cause population growth commensurate with economic opportunity. Increase through the birth rate produces a more homogeneous people and obviates many of the social and political problems which are created by the process of expansion through immigration of aliens.

Many important consequences will flow from the policy of restriction, if this policy is applied for any considerable period of time. Of chief importance perhaps will be its effect upon different labor groups in the country and upon relative wage rates. Native-born Americans rarely become unskilled laborers. A drastic reduction of the relative supply of workmen in this group must result from the new immigration policy. This will raise the lowest rates of

wages in the country and tend to reduce existing differences in wage rates. Restriction of immigration will also promote transference among the different groups of workers by allowing the system of free education through trade schools, night schools, and the like to reach almost the entire laboring population. Formerly the workmen of alien birth — an important fraction of the whole — were but little influenced by the educational system. This too will promote mobility of labor and work toward greater uniformity of earnings.

The restrictive policy furthermore will have its effect on the organized labor movement. One great obstacle to the spread of this movement heretofore has been the heterogeneity of the population. The native-born skilled craftsmen have shown little tolerance for or sympathy with the immigrant workers; they have made no sincere effort to bring the latter within the existing unions or to aid them in forming unions of their own. As a result trade unions have flourished almost exclusively among the skilled and have shown the narrowness of outlook and selfishness of purpose to be expected from a movement thus limited in its membership. These conditions will change as, with the decline of immigration, the laboring population grows more homogeneous in language, customs, and economic prosperity. The probable effect will be to strengthen the labor movement and profoundly to affect its policy in the direction of a broader and more constructive outlook. The changed conditions will be most noticeable in those branches of industry — *e.g.*, textiles, the mines, the steel industry — which employ unskilled workers and which in the past have been able to rely upon an abundant supply of this type of labor. The increased bargaining power which results from relative scarcity of supply will be reflected not only in the wage rate but in many other phases of the relationship of the workers with their employers.

XLIX

UNEMPLOYMENT

The meaning of unemployment. The unemployed, as the term is used here, consist of workers unable to find employment though seeking it and competent to perform work. This definition, it will be seen, does not include all who may be idle at a given time. Those who won't work — wage earners on strike and the work-shy — are not included in our definition. They do not in fact constitute more than a trifling part of the total unemployment. It is a more difficult problem to decide whether the definition of unemployment should include the aged and the disabled — people who have been wage earners and are now seeking jobs, but who, through accident or disease or because of the preference given to younger workers, have lost their ability to find employers. For the most part these unfortunates are victims of our modern system of production. Their condition is frequently the result of the health hazards of industry or of the tension and strain of present-day productive methods, which place a premium on youth. The statistics which we shall use in estimating the amount of unemployment include this group of people; but when dealing with the causes and remedies of unemployment it is well to keep them separate from the rest of the involuntarily idle, since in these respects the aged and disabled present a problem peculiar to themselves.

Extent of unemployment. Attempts to estimate the amount of unemployment in the United States have been handicapped by the absence of comprehensive and accurate statistics. Trade unions have kept records showing the amount of involuntary idleness among their members; governmental bodies and research organizations have examined the payrolls of industry as a basis for measuring the rise and fall of employment; censuses of the unemployed have been taken in selected cities and at least once (1930) throughout the country as a whole. These studies have sufficed to make it clear that the number of unemployed is always

large, and astonishingly large in times of general business depression; but they have not given us a total figure over a series of years which we can accept with confidence in its accuracy.

The need for reliable statistics has become urgent now that the United States has set up a program of unemployment insurance. Interest in this aspect of the problem and the attempts to relieve the widespread distress of the recent years of depression have given renewed impetus to the study of the extent of unemployment. We shall draw upon these recent statistical computations for the purpose of our inquiry. The figures given are the best available, the product of careful and prolonged study. At best however they are estimates, not the result of actual count, and as such they probably contain a considerable margin of error.

The table given below shows the percentage of the combined labor supply of the major American industries unemployed in each of the years from 1900 to 1920. These figures, it should be noted, include the sick and disabled, who would be excluded by a strict application of our definition of unemployment. Workers in agriculture, in mercantile establishments, and in the public utilities are not included.

Unemployment in Manufacturing, Transportation, Building Trades, and Mining, 1900–1920 [1]

Year	*Percentage unemployed*	*Year*	*Percentage unemployed*	*Year*	*Percentage unemployed*
1900	10 0	1907	6 9	1914	16 4
1901	7 5	1908	16 4	1915	15 5
1902	6 8	1909	8 9	1916	6 3
1903	7 0	1910	7.2	1917	6 0
1904	10.1	1911	9 4	1918	5 5
1905	6.7	1912	7 0	1919	6.9
1906	5.9	1913	8 2	1920	7 2

Taking a view of the period as a whole, the table shows that approximately 10 per cent of all workers attached to these four groups of industries were unemployed on the average. In the best

[1] P. H. Douglas and A. Director, *The Problems of Unemployment*, 1931, p. 28.

years unemployment remained at about 6 per cent of the working force; in years of depression it rose abruptly. The figures would be somewhat reduced by the inclusion of the relatively more stable occupations omitted from the table — public utilities, trade, domestic and professional service. The investigators have estimated that to take account of these occupations and to exclude the aged and disabled would reduce the percentages by 2 per cent, indicating that the average true unemployment for all industry has been approximately 8 per cent. It is observable that the volume of unemployment has shown little tendency either to increase or diminish during this long period, the major fluctuations from year to year being attributable to conditions produced by the business cycle or by war.[1] These estimates of chronic unemployment have been given in percentages of the labor force rather than in absolute numbers, because of the non-inclusive character of the group covered by the table. At no time during the period did the actual number of the unemployed in these four industries fall as low as 500,000 people.

The table on the following page shows the course of unemployment since 1920, both in actual numbers and in percentages of the wage-earning population. These estimates cover all types of employment except agricultural workers.

It will be seen that even in 1926, which was a fairly good year industrially, upwards of a million and a half workers were unemployed. The disastrous unemployment which accompanies a prolonged depression is reflected in the figures following 1930. These figures, large as they are, do not tell the whole story, for the well-established practice of spreading work by placing the whole labor force on part time has reduced the percentage of complete unemployment while throwing a burden of partial idleness upon the workers who still hold jobs. In March, 1932, the President's Organization on Unemployment Relief made a survey of the employment situation in 6,551 manufacturing concerns, whose total labor force in normal times numbered some 3,500,000 wage earners. It was found that not only had 27 per cent of these workers lost their jobs, but that over half of those still employed

[1] Douglas and Director, *op. cit.*, pp. 32–33.

were losing nearly 40 per cent of their earnings through part-time employment.

Estimated Unemployment in the United States [1]

Year	*Average monthly number unemployed*	*Average percentage of all wage earners*
1920	1,400,000	5.1
1921	4,300,000	15.3
1922	3,400,000	12.1
1923	1,500,000	5.2
1924	2,300,000	7.7
1925	1,800,000	5.7
1926	1,700,000	5.2
1927	2,100,000	6.3
1929	1,100,000	2.3
1930	4,100,000	8.4
1931	8,200,000	16.6
1932	12,700,000	25.4
1933	12,900,000	25.4
1934	11,200,000	21.8
1935	10,300,000	19.9
1936	8,600,000	16.4
1937	7,400,000	13.7

Causes of unemployment: Demand and supply of labor. It is obvious that unemployment is always a symptom of maladjustment between the supply of labor and the demand for labor in the different occupational groups and industries. Assuming as a starting point a condition of equilibrium between demand and supply in the labor market, a search for the causes of subsequent unemployment will lead us into a study of changes which have occurred on one side of the market or the other. Now there is one general characteristic of the supply of labor which, though not a positive cause of unemployment, does have an important influence on the amount and the duration of idleness resulting from other causes. This is the relative immobility of labor, which impedes the transfer of workers from place to place and from one kind of occupation to another in response to changes in the conditions of demand.

[1] Figures to 1927 from L Wolman, *Recent Economic Changes*, Vol. II, 1930, p 478 Later figures based on studies of the National Industrial Conference Board The year 1928 is not included by either of these sources

Aside from the above mentioned supply factor, the causes of unemployment are the recurring unforeseen changes in the demand for labor inherent in the unstable and dynamic character of our industrial system. Among these changes are some of relatively minor importance, such as the sudden and capricious alterations of fashion in dress and the ephemeral fads of amusement and entertainment which are so representative of our volatile age. Other changes of a sporadic nature are caused by such events as the regional migration of industries, the reorganization of formerly independent enterprises into larger single units, and the like. All of these fluctuations tend to upset the equilibrium of the labor market, but their effects are so slight in comparison with those of the major forces of change that we shall leave them out of account. The major causes of unemployment may be classified in three general groups: (1) technological change; (2) seasonal fluctuation of industry; (3) the business cycle.

Technological change. Improvements in the technical processes of industry are inspired by the constant pressure to reduce the cost of production. They take many and varied forms: the invention of new and the bettering of old machine equipment, the development of more efficient organization of labor and machines within the factory, the improvement of methods of management. In the case of most of these improvements, the saving takes the form of a reduction in the wages cost of producing a given quantity of goods, either by diminishing the number of workers employed in the process or by substituting a cheaper for a higher grade of labor. Looking back through history, examples can be found of sweeping alterations in certain occupations produced by invention. Of such significance was the invention of the power loom, the nail-making and bottle-blowing machines, the linotype machine. The period since 1919 provides the most striking illustration of technological advance because, as is shown by reliable records, there was a rapid increase in the total product of industry, generally accompanied by a large decline in the numbers of men employed. This of course was due to the introduction of labor-saving machines. Between 1920 and 1929 the output per unit of labor (labor-hour) in all lines of production rose 25 per cent; in manufacturing alone

the increase was 45 per cent. This meant of course that the 1920 volume of output could have been maintained ten years later with a labor force reduced by these percentages.

All this does not mean however that the demand for labor was in fact reduced by these technological improvements. Vastly more labor was required in the making of cotton and woollen textiles, nails and bottles, and in the printing trade, after the introduction of the epoch-making inventions mentioned above than was employed in the earlier hand-work stage. Coming down to later times, the period 1919 to 1929, which was marked by such extraordinary introductions of labor-saving machinery, was the very period in which we find the most striking decrease in unemployment, as shown in the table on a previous page of this chapter. The reason for this should be clear to the student of economic theory. The effect of such technological improvements is to lower the cost, and therefore the price, of the good in question. A greater quantity of the good will be sold at the reduced price; the industry will expand, thus tending to reabsorb the temporarily displaced workers. For example the 1920 volume of output of American industry was not sufficient for the people's needs ten years later. The actual volume was so much increased that industry more than absorbed the number of workers apparently displaced by machines.

The reader's knowledge of value theory will make it clear that the amount of expansion and reabsorption of labor which results from price decline in the particular industry affected will depend on the elasticity of the demand for the good. If the elasticity is greater than unity, it is possible that even more labor will be employed in the same industry after the change than before; if less than unity, the labor force in that particular industry may decline. But in this case the consumers of the good will have spent a smaller total amount of money in satisfying this particular want, and the portion of social income thus released will be spent on other goods and services. The net effect therefore will be to reduce the demand for labor in one line of production and to increase the demand in others. This result will follow regardless of the manner in which the saving is distributed within the social group. If the owners of the industry are enabled by monopolistic practices to keep all the gain

to themselves, they will have additional incomes to spend. If those who gain income from the new process — whether owners or consumers — elect to invest rather than to spend their gains, this too will create a demand for labor; for what is invested is ultimately spent. These various alternatives will determine whether the new demand for labor takes one form or another, but they will not affect the amount of this demand.

From the social point of view therefore the ultimate effect of general technological improvement upon employment is greatly to increase the demand for labor while redistributing the labor supply. If the changes are very rapid and widespread, as in the decade following the World War, the shifting of the demand for labor may take place on a wide scale and give rise to entirely new types of occupation. Needless to say, society gains from this process in more abundant and more varied consumable income, and the working classes share in this gain along with everyone else.

But this general and long-run view of the matter does not entirely dispose of the problem of technological change as it affects the individual worker. He is interested in the effect of changes in his industry upon himself, and in their immediate effect. If he is deprived of his present job, he cannot transfer immediately and without loss to another locality; and if he must abandon his present craft or skill, he may be unable to take up another save on a lower level of wages. When changes are occurring on a large scale, the new demands for labor do not spring up simultaneously with the decline of the former demand; and when these new opportunities for employment do appear, the immobility of labor prevents the displaced workmen from taking immediate advantage of them. Hence technological change causes temporary unemployment, temporary in the long view of things, but serious or even disastrous to the workers affected by it. Comparing the employment records of 1929 with those of 1919, we find 500,000 fewer workers employed in manufacturing in the United States; 300,000 fewer on the railways; 200,000 fewer in the mines; while of the 3,800,000 people who left the farms in this decade, at least 1,500,000 were of working age. In the main these displacements of labor were the result of improvements in production methods within the industries con-

cerned or — as in the case of the railroads — to technological advances in competing industries. During the prosperous years 1924–1927, the unemployed in this country averaged nearly 2,000,000 people.[1] Numerous case studies were made of the job-hunting experiences of displaced workers during this period. While we cannot review the results of these inquiries in detail here, it is significant that a large fraction of the workers studied were unemployed for over a year; many who found jobs suffered permanent reduction of wage income; and a considerable number had found no regular employment at all, though we should still be reminded that during this period total unemployment was decreasing.

Society as a whole undoubtedly benefits from technological progress; for this reason it is unwise and short-sighted to oppose improvement in production methods. But it is unfair that this general benefit should be had at the cost of some members or classes in society. The logic of the situation is to the effect that the costs of unemployment caused by technological progress should be borne by society as a whole and not thrown solely on those who have the misfortune to feel their effects in the first instance.

Seasonal fluctuations: Causes. The forces which cause seasonal unemployment are so diverse and conform so closely to the peculiarities of the individual enterprise which they affect that they must be studied in each case as technical phases of particular industries. In general however their effects appear in the form of an unequal distribution throughout the year either of the demand for the product of the enterprise or of the supply of raw material which it utilizes. For example in the clothing trade there is a sharp concentration of demand for certain types of clothing in brief periods of the year. The demand for the product of coal mines varies markedly with the seasons. The erection of buildings is seasonally affected both by changes in the demand and by climatic conditions which limit operations. The manufacture of foods, such as canned fruits and vegetables, fluctuates in sympathy with the seasonal nature of agriculture. These are some of the more visible effects of this factor, but its influence is widespread. Some enter-

[1] Leo Wolman, *Recent Economic Changes*, Vol II, 1930, p 478.

prises are affected indirectly by reason of their position as subsidiaries to other industries whose market is disturbed by seasonality of demand; an example is the cement industry, which feels the effects of climatic control of building operations. Manufacturers of specialties, who sell a finished product to the ultimate consumer, find that the demand for their output fluctuates as a consequence of changes in fashion or in harmony with the forces of habit and custom.

The following type cases will indicate the importance of this factor as a cause of temporary unemployment. In the coal industry 40 per cent more workers are employed in the most active as compared with the least active month. The construction industry in Ohio during normal years employs on the average 79 per cent more labor in September than in February. An index of employment in retail trade in 1929 shows a variation of 30 per cent between the best and the worst months. A similar index for the clothing trade showed a fluctuation of 35 per cent, and for automobile factories, a fluctuation of 24 per cent.

The labor reserve. It is the tendency of industries subject to wide seasonal variations to attract to themselves a disproportionate number of workers as compared with the more stable enterprises. The attractiveness of the relatively high wage rate, amplified by bonuses, overtime pay, and other devices upon which employers rely to stimulate production during the season of insistent demand, causes the workers to lose sight of the obvious fact that their annual earnings must be reduced by periods of enforced idleness. The supply of labor permanently attached to such industries tends to be equated to the highest seasonal demand. In other words, there is a *reserve of labor* in these enterprises, an excess of workers over and above the number who could be continuously employed if the seasonal variations were removed. This labor reserve is drawn upon to equip the industries during the period of intense activity and is refilled by the discharge of workers as the volume of sales diminishes and business grows dull.

Consequences of seasonal unemployment. The social injury which flows from this state of affairs is the loss of labor power, as measured by the average idleness of the working force. This loss

of work must be charged as one of the costs of operating the industry, regardless of the fact that the employers bear a wage expense equivalent to only a part of the social loss. The fact that the unemployment is shared piecemeal by the entire labor group in no wise qualifies the principle involved, though it does conceal the loss and may act as an obstacle to its correction.

The effect upon the laborers in the seasonal industries is twofold. In the first place, it makes the actual earnings of the worker less than would be indicated by the wage rate. So clear is the first of these effects that it could be dismissed without mention except for the abundant evidence that it is imperfectly understood, both by the general public and by the wage earners themselves. Whenever a strike occurs in a seasonal industry, such as coal mining or the building trades, certain persons seize on the relatively high daily rate of wages as evidence that the workers are presenting exorbitant demands. It is also to be observed that the incoming generation of workers, impressed by the higher daily rates in seasonal as compared with stable industries, neglects to discount the higher rates by translating them into normal yearly earnings. It is for this reason that seasonal industries tend to build up a labor reserve.

Quite apart from its effect upon the amount of earnings, persistent recurrence of unemployment reduces the utility of the wage earner's income. The irregular distribution of earnings throughout the year is equivalent to withdrawing wages from a period when their marginal utility is high and returning them at a time when their marginal utility is relatively low.

The business cycle. The different phases of the business cycle have diverse effects on the welfare of labor. During the prosperity period there is a tendency for wages to rise less rapidly than other prices, with the result that the real income of the wage earners falls. Nevertheless it is probably true that the net effect upon the workers of this phase of the cycle is good. This is a period when unemployment is reduced to a minimum. Members of the family not usually employed are drafted into service. Overtime pay at rates higher than the ruling scale of wages, bonuses, and other perquisites are bestowed upon the worker by the employer in his effort to maintain a stable working force.

The problem of unemployment appears at the time of the crisis. This condition is found not only in one enterprise or group of enterprises but in all types of industries throughout the entire country. Consequently the unemployed can find no escape from the disaster in other industries or other places. They have no alternative but to remain idle and attempt to tide their families over the "hard times" by drawing upon the slender reserves accumulated during the period of prosperity. The savings of years are sacrificed, household effects are surrendered to the pawnbroker, homes are mortgaged. Experience of these consequences of business depression in America has been too recent and too painful to require much elaboration of the topic. When conditions were at their worst following the collapse of 1929, a fourth of the wage earners of the United States were unemployed, millions had been forced into pauperism, charitable organizations had exhausted their ability to cope with the problem, and the combined resources of city, state, and federal governments were strained to keep the people from starving. At such times many people are driven to crime. There is widespread unrest among the sufferers, leading to organized movements of revolt and revolution.

As we have learned from previous study of the business cycle, the period of depression is one in which the price system and the industrial structure are seeking a new point of equilibrium. Eventually this point is found, and business activity begins gradually to quicken; but a long and painful period of readjustment may be required to restore the balance. For the working population readjustment requires that they shift in large numbers to new localities and new occupations, so that even when equilibrium is restored many of them will have suffered irreparable losses through the severance of home and community ties and the abandonment of their former crafts or trades. These losses must be added to the direct costs borne during the period of actual unemployment.

Attitude of the laborers. The fear of unemployment exerts a powerful influence upon the behavior of the wage-earning classes of society. Its causes are for the most part beyond their control. Yet it is they who feel its effects most grievously, and it is but

natural that they should seek to protect themselves by using any weapons which lie in their power. When labor is organized in defense of its interests, the union is prone to adopt policies designed to prolong the job or to spread it over a larger number of workers. One such policy of union labor is the demand for what is known as the "full crew." Agreements between trade unions and employers in the building trades often specify that each group of craftsmen be accompanied by a minimum number of apprentices and helpers. Labor organizations on the railroads have demanded that each train be staffed by a certain number of workers of various classes without much regard to the real utility of the full crew in the circumstances of the case. The seamen's unions have advanced similar demands with respect to ocean-going vessels Other examples of this line of trade union policy could be given They are in harmony with the wage earner's habit of looking upon productive enterprises as a job-creating agency.

But the most far-reaching effect of the fear of unemployment upon the conduct of the workers is seen in the disposition of labor — the unorganized as well as the members of trade unions — to oppose measures designed to increase production In earlier days the workers frequently resorted to violent methods to destroy the new machine, nowadays they are more likely to cripple its functioning by concealed sabotage. It is impossible to calculate how much society loses through the time-wasting devices of workers in fear of unemployment. Make-work tactics are not founded on any closely reasoned economic policy, but they do reflect a general belief that there is at any given time only a certain amount of work to be done. The problem of each laborer is to get his share of this quantum of employment. It is assumed that if one man through extra exertion does the work of two ordinary workers, he drives someone else into unemployment, if all workers discharge functions with unusual efficiency, they will all be forced into idleness a part of the time; employment is a thing to be husbanded, economized, made to last.

It is not difficult to prove that make-work tactics are injurious to social welfare A general indulgence in inefficiency can lead only to general impoverishment. But this does not prove that make-work policies are always unwise from the standpoint of the

particular men who practice them. Where they have the power to control the supply of their type of labor, a group of workers may increase the regularity of their employment by decreasing their rate of output. On the other hand, it is sometimes true that an increase in efficiency on the part of a particular group of workers will either force some of them into other lines of employment or reduce their wage rate. When no collective attempt is made by society to solve the problem of unemployment or to lighten its effects upon the wage earners, it is not to be expected that any specific labor group will refrain from adopting measures effective for their own protection merely because these measures react injuriously on other social classes and other groups of workers.

Fear of unemployment will explain other practices of the trade unions which are not always understood by those outside the wage-earning group. The importance of the job to the worker is so great that he develops a proprietary attitude toward it whenever conditions are favorable. Strongly organized groups which have entered into collective agreements with their employers fall into the habit of viewing the job as the joint property of employer and worker. They believe that neither party to the labor contract has the right to dispose of these jobs without consulting the other party. An understanding of this view of the situation throws light upon the behavior of men when on strike. As viewed by the law of the land and by that part of the general public which shares the employers' philosophy of property rights, a strike is a voluntary surrender of employment by a group of workers acting collectively. This action has freed the employer from any obligation toward his former labor force and has authorized him to offer employment to anyone in search of work. But the strikers do not acquiesce in this opinion. They have quit work with the intent of retaining the job; the strike is merely a contest over the terms of employment between the two parties who jointly control employment in the enterprise. Until this contest is settled, the enterpriser does not have a free hand in filling the vacant places. Any worker who accepts employment during the strike is a job stealer, a "scab"; and against him aggressive measures may be employed with propriety. Here we have a clash of ideas so fundamental that it is

sometimes impossible to reconcile the two points of view without use of force. Hence labor disputes often result in violent conflict between the strikers and the men who accept employment in the enterprise.

Remedies for unemployment: Adjustment of labor supply. There are two ways of attacking the problem of unemployment. One is to promote readier adjustment of the supply of labor to the changes of demand which result from the instability of modern industry. The other is to attempt to reduce the fluctuations of economic enterprise which through changes in demand cause the volume of employment to rise and fall.

Considering the first of these methods, we may again refer to the immobility of labor as a factor of major importance in the maladjustment of supply to demand. This appears in two different forms, *functional* and *geographic immobility*. From what has been said of the causes of unemployment, it is clear that some of the trouble is caused by the inability of the laboring population to shift from one locality to another or from one calling to another in response to changed conditions. It follows that any policy which promotes the versatility of the average wage earner will act as a correction of unemployment and that any device which increases his geographic mobility will have the same effect.

Versatility of labor: Education. Acquisition of skill is a matter of education. Since our productive system in the past has required a high degree of specialization, the educational process both in the schools and in apprenticeship has been designed to give young people of the wage-earning class a single type of skill. This practice is based on the assumption that the typical worker will follow the same trade throughout his life, an assumption which has been partly invalidated by the recent rapid pace of technological change. Society has not yet devised a system of vocational training adapted to its own needs. This situation would be in part relieved if education of the trade schools were of a broader and more flexible kind than in the past and specialized training in a given craft were accompanied by supplementary training in other lines. Much can be done as well in the re-training of adult workers displaced from their regular employments and making opportunities to acquire

new skills generally available to such workers at public expense. It does not fall within our province to discuss the details of the educational programs required to accomplish these purposes, but their economic importance is beyond dispute.

Geographic mobility: Private employment offices. Geographic mobility of labor can be increased and unemployment reduced by increasing the worker's knowledge of available jobs in his own or another locality. This is the function of the employment office, sometimes known as the labor exchange. A complete system of employment offices would cover the whole industrial area with an intercommunicating series of local agencies, each registering and classifying the idle workmen and the available jobs in its locality and all coöperating to distribute unemployed labor over the entire area with as little delay as possible. In the absence of such a system publicly maintained, the service will be supplied as a profit-making enterprise by private initiative.

The private employment office however has been found to develop into an evil unless controlled. The owner has no interest in reducing the amount of unemployment; his income is drawn from fees charged the idle workmen who register at his agency. Instances have been known of collusion between the owners of labor exchanges and the foremen or employment managers of industries; the foremen agree to discharge all newly employed workmen as speedily as possible so that a steady stream of applicants may pass through the exchange on the way to the job. Many times the fees charged for the service of the exchange have been exorbitant, especially during periods of acute distress. Private exchanges have also sent workmen on futile journeys to distant cities equipped with cards of introduction to fictitious employers. They have worked in collusion with loan sharks who advanced money required for railroad fares at usurious rates on chattel mortgages. Experience with this institution has proved that it is unsafe to entrust the solution of a social maladjustment to a private enterpriser who draws his profit from its continuance. And even if these abuses do not develop, the private employment office is incompetent to meet the need for systematic and unified service over a wide area.

Public employment offices. The only effective solution of the problem is for the government itself to assume the function of supplying idle workmen with information regarding opportunities for employment. Great Britain has a system of publicly owned and managed employment exchanges covering the entire country and giving its services free of charge to the workmen. This institution not only serves to distribute the labor supply during periods of unemployment but is a valuable agency for the collection of employment statistics to be used as the basis for a more thoroughgoing solution of the problem. Other countries too have well-organized systems providing efficient service on a nation-wide scale. In America this obvious corrective of unemployment has only just now begun to be applied in a thoroughgoing way. Many of the states have maintained public employment offices; but with few exceptions these have proved virtually impotent to discharge their function because of insufficient financial provision or for less creditable reasons.

The federal government has provided some employment service since 1907, but except for the war years the service has been of trifling importance. A fresh beginning was made through the Wagner-Peyser Act of 1933. This created the present United States Employment Service with a central office in Washington, staffed with trained people under a director-general, who has the assistance of an advisory council representing employers and workers. To stimulate state action the law empowered the federal government to make an annual appropriation of $4,000,000, of which $3,000,000 is allotted to the states according to population, on condition that each state which accepts the grant shall match the federal subsidy dollar for dollar and that the state's system of employment offices shall measure up to standards defined by the federal service. The remaining $1,000,000 is used to conduct an interstate clearing house of information, to inspect the state services, to carry on research and publication, and to render certain direct services to labor in the District of Columbia and among the war veterans. At first the states were slow to respond to the invitation of the federal government to revamp their employment offices and affiliate with the United States Employment Service. By June, 1935,

only 25 states had done so. But the passage of the Social Security Act in 1935 hastened the process. This law embraces a broad group of policies which we examine in another place; but we must note here that its provision of federal aid to state unemployment compensation systems by necessitating the creation of state-wide employment offices to record the unemployed and distribute the compensation spread the public employment office system throughout the entire country. By 1938 all the states and territories had set up these systems and all were affiliated with the United States Employment Service; at this time there were 2,000 offices in the system, employing a staff of 15,000 people.

During the brief period of its existence the Service has been to a large extent deflected from its normal long-range functions to discharge placement services required of it by various federal emergency policies such as the CWA and the PWA. Its record nevertheless is impressive. From 1933 to 1938 some 26 million unemployed people were registered in the employment offices of the country as applicants for jobs; about 22 million placements were made, including those in the various work relief services. Toward the end of this period an increasing proportion of placements was being made in private industry.

Stabilizing the demand for labor. After everything has been done to facilitate the adjustment of labor supply to changes in demand, the problem of unemployment will remain to a large extent unsolved. The ultimate causes of the problem are to be found in the forces of change which produce sporadic, seasonal, and cyclical fluctuations in the demand for labor. Attempts to deal with these forces directly have called forth a wide range of proposals, all having as their purpose a partial or complete regularization of the activities of modern business enterprise. One group of these proposals deals with the problem of the business cycle. Since this phase of the subject has been examined in another place in this book, we shall not consider it here further than to point out the obvious fact that any policy which mitigates the extreme swings of the business cycle must have an effect on the problem of unemployment. The other policies which have been suggested for the stabilization of the demand for labor may be classified into two

groups: (1) those which rely upon voluntary action by employers and (2) those which must be administered by governmental authority.

Stabilization by private enterprise. Within recent years managers of enterprises in seasonal industries have taken steps to stabilize their demand for labor on their own initiative. The motive behind these attempts to reduce unemployment has been a desire to increase efficiency of labor and reduce overhead costs. Modern employers have learned that the constant fear of unemployment is a potent cause of the tendency of wage earners to restrict output. Furthermore the turnover of labor in a given enterprise involves expense of hiring and training the workers which could be reduced by maintaining the labor forces at a constant level and retaining the same workers on the payroll for a considerable period of time. In the clothing markets of Chicago and Cleveland it has been found that the wage earners are quite willing to guarantee a standard rate of production, provided the employers in turn guarantee continuous, or nearly continuous, employment. The experience of many large concerns has also proved that wage earners will attach themselves permanently to that employer who can promise a lifetime job.

A detailed study of the devices by which seasonal industries are stabilized lies in the field of business management and outside the limits of this book. The following is a summary of the more important of these devices. (1) An industry whose major product is subject to fluctuating consumers' demand may take on *compensating side lines* whose markets vary inversely with that of the major product. An example is the manufacture of felt hats or cloth caps by a concern specializing in straw hats, or the production of some staple product by a concern devoted primarily to the manufacture of Christmas specialties. (2) Somewhat similar is the development of *compensating markets*. Cement manufacturers have regularized their output by exploiting the South American markets where the seasons are the reverse of those at home. Much can be done on a smaller scale along this line within the continental area of the United States. (3) Goods may be *made to stock* during the slack months. This device involves expense of storage and can be used

with safety only in connection with types of product which do not deteriorate in storage and are not subject to capricious changes of fashion. (4) Sales methods may be adjusted to encourage *advance orders* with a long delivery time. By offering particularly favorable terms, a large concern specializing on the Christmas trade has succeeded in inducing many of its customers to place their orders more than a year in advance, thus allowing a stabilizing of production. Similar results have been obtained by many of the larger clothing manufacturers. (5) Companies which perform several different processes may set up a system of *vocational training,* increasing the versatility of their workers so that they may be transferred from slack to busy departments without loss of time.

When a whole branch of industry is organized in a coöperative trade association, these and other methods of stabilization become more effective. In America organizations of this sort have been handicapped by the danger of conflict with the laws forbidding restraint of trade, yet there have been some successful experiments along this line, particularly in the clothing trades. A still wider extension of effort to stabilize employment, involving coöperation between an employers' association and organized labor, has been attempted at various times. These union-management agreements have proved effective in preventing abrupt changes in demand for labor due to technological advances. Thus the introduction of the linotype machine into the printing industry was controlled by an agreement between employers and typesetters which regulated the rate at which the new method was adopted, allowing time to train the former craftsmen to operate the machine. The union of hand moulders, after struggling vainly to prevent the adoption of machine methods in their craft, made a similar agreement with their employers for joint control over the rate of technological change. Very recently the textile workers and the national association of their employers have attempted by joint agreement to provide for more gradual adoption of the improved automatic loom.

Stabilization by governmental action: In general. In every period of widespread unemployment there arises the demand that the state take over the task of solving the problem. Extreme pro-

posals involve thoroughgoing revolution of the economic system, abolishing private enterprise and competition, and substituting a socialistic organization controlled throughout by governmental authority. The programs of planned economy which propose to schedule the rate of production in each branch of economic activity and to keep the whole system in perpetual balance by centralized regulation are of this type. At a later point in our study we shall examine these programs of socialism. Here we are concerned only with such policies as can be applied within the framework of the existing economic order.

Proposals for governmental action have relied for the most part upon the spending power of the modern state. The more moderate of these policies propose merely that the government manage its own routine employment of labor in such a way as to increase demand during periods when private enterprise is at a low ebb. It is conceivable that the routine construction work and other such normal activities of local and central governments might be concentrated on periods of seasonal unemployment In fact this experiment has been tried with some beneficial effect. But these activities of government usually create a demand for skilled labor of a specialized type — for example, labor of the building trades — and not for labor in general. The machine operative, the clerical worker, the unskilled laborer do not find much opportunity to transfer for temporary employment to the public service, however carefully the government may have scheduled its routine program.

Public works: Theoretical analysis. The public works program in its broader form is a different matter. This contemplates governmental activities on a large scale during periods of general business depression, involving works which would not otherwise be undertaken at all, such as slum clearance, reforestation, road building, irrigation projects. The idea in its theoretical form calls not only for advance planning of a comprehensive public works program but for careful training of expenditures to correlate with the "troughs" in the fluctuations of private business.

As advanced by serious students of the problem, the public works program is to be safeguarded against the most obvious dangers involved in extravagant governmental spending. The program is to

consist entirely of projects worth while for their own sake; it is to be financed as much as possible from funds reserved for the purpose during prosperous years, thus avoiding the necessity of heavy taxation or large public borrowing during a time of unsettled business; it is to be administered in a businesslike way, employing laborers on the basis of their usefulness to the enterprise, not merely because they are unemployed, and holding them to the same standards of wages, working conditions, and performance demanded by efficient private enterprise. In addition to these safeguards, the public works program is to be based on reliable statistics of unemployment, to permit of advance planning in prosperous years and to regulate the scope and the regional location of the government's activity in harmony with the trend of unemployment. If all these features of the program can be assured, there is much to be said for such a program as a corrective of cyclical unemployment. It proposes to do two things at the same time: prevent distress among the unemployed, and add to the nation's equipment of socially desirable capital produced by labor which would otherwise have gone to waste.

These theoretical considerations place the public works program in an attractive light, but misgivings of a practical nature arise when such a policy is pictured in actual operation. An essential feature of the plan is the accumulation of reserves during prosperity years. But what is to be done with these reserves? Are they to lie idle in the form of money until needed, thus disturbing the effective money supply of the country and consequently the price level? Are they to be invested in income-bearing securities? If so, will this not disrupt the security market; and how in a depression when the market is prostrate can the reserves be turned into cash without disastrous result?

In addition to the serious problem of finance, there is an almost insuperable problem of administration during the active phase of public works construction. But let us assume for the moment that wise administration of the public works program is assured; there would still be reason to fear that its indirect effect on the situation might prove injurious. The ultimate solution of business depression must be a readjustment within the price structure and cor-

relatively within the system of production which will bring productive forces to a new point of equilibrium. This final objective can be achieved only if individual prices are free to move in harmony with changing conditions and the supplies of labor and capital in the different lines of activity are also free to respond. Anything which impedes these changes is likely to prolong the maladjustment and hence to postpone final solution of the problem. Now a gigantic program of public works cannot fail to create an artificial demand for certain types of goods and labor, which will tend to maintain existing prices and wage rates in the industries affected by the program. Obviously there is danger here that the government's well-intentioned efforts to ease labor's burden of unemployment will delay fundamental solution of the problem to the detriment of society in general. This sort of indirect and unforeseen effect of governmental intervention in economic affairs it is virtually impossible to prevent, even by the most careful administration.

Returning to the question of administration, there is grave doubt lest this phase of the program go astray. Experience of governmental activities in economic matters emphasizes the dangers of extravagance and political chicanery inherent in any policy of widespread spending. At best it is doubtful whether the difficult task of advance planning can be discharged in an impartial and scientific way by governmental agencies, while there is always the chance that political pressure at the time the program is put into effect will disrupt these plans and cause expenditure upon projects which are sheer waste, or worse. Businesslike management of labor by public officials dependent on political support must also prove extremely difficult to accomplish. A spending program of the dimensions contemplated will give the political party in power at the moment such opportunities to consolidate its influence throughout the country that it is questionable whether professional politicians can resist the temptation to use this power for partisan purposes. Misgivings on these points arise in the minds of cautious people who are quite ready to admit that, in the abstract, this proposal for the solution of the problem of unemployment possesses merit. It goes without saying that if the public works program falls victim of these political hazards, it will do no good and may make a

bad situation worse. Waste of public money on projects which serve no useful purpose will not create additional employment but merely expend on objectless labor funds taken from the nation's citizens.

Public works program of the United States. The merits of this ideal public works program have never been put to the test of experience, for nowhere in the world has such a program been attempted. It is true that a vast program of public spending has been undertaken by the United States government as a remedy for unemployment, beginning with the National Industrial Recovery Act of 1933, which appropriated three and a third billion dollars for this purpose. This and subsequent appropriations extending through 1938 have been spent, in part as subsidies to state and local governments to assist in financing approved projects designed and administered by these governments, and in part on a number of large-scale enterprises undertaken by the federal government itself. The program is under the immediate direction of the President and the Secretary of the Interior, who must approve all projects proposed by the state and municipal authorities and see to it, as a condition of federal subvention, that these projects receive equivalent financial support from the local governmental units. This experiment has undoubtedly done much for the relief of distress and may have exerted some influence in the business recovery of 1934–1937. But whatever judgment is passed upon it, it does not constitute a real test of the public works method of solving unemployment. This particular program was adopted as a substitute for an existing policy of unemployment relief which was draining the resources of governments, local, state, and national. It had no such background of long-range planning as is proposed as a fundamental feature of the real public works program. For these reasons the labor employed is selected or managed, not by tests of efficiency, but as objects of charitable relief; and the projects undertaken have been devised on the spur of the moment in the midst of a great national emergency. Moreover the program was financed, not out of reserves accumulated over previous years, but by current taxation and borrowing at a time when the investment market was already disturbed and the taxpayers already

greatly impoverished by the prolonged business depression. In all these important respects the present activity of the government fails to qualify as a true test of long-range planning of public works.

Unemployment insurance: Practicability. There is little hope that devices for adjusting either the demand or the supply of labor will prevent unemployment entirely. The major fluctuations of the business cycle are as yet beyond human control, and changes in technical process and in consumers' demand will doubtless continue to upset the labor market. It must be assumed that a certain amount of unemployment is an inevitable feature of our system of industry. In the case of other ineradicable hazards, systems of insurance have been developed to distribute the losses and reduce the burdens borne by the immediate sufferers. Cannot the same method be applied to unemployment?

At first sight unemployment seems to conform to the tests of an insurable risk. It applies to a sufficiently large group of people and over a sufficiently diversified range of conditions to permit of the operation of the law of averages. By statistical study the normal rate of unemployment for each of the larger groups of industries can be determined with sufficient accuracy to make it possible to predict the number of unemployed and to provide in advance for their relief. The economic loss involved in employment is definite and ascertainable; it can be wholly or partially compensated by indemnifying the idle workman on the basis of his former earnings.

Not only do these features of the unemployment hazard appear to make it amenable to the principle of insurance as a practical matter of business, but weighty arguments of a general social nature can be urged in favor of the policy. On grounds of justice it appears unfair that those least capable of bearing the burden and least responsible for its causes should be called upon to suffer. Furthermore recognition of the fact that the risk of unemployment is a cause of far-reaching productive inefficiency among the workers argues that society would gain more in increased productivity from a policy guaranteeing the security of labor than it would lose through the expense of administering the policy. Finally the only real improvement in the situation consists in preventing

unemployment or reducing its amount, not in expedients for distributing the loss occasioned by it. Little progress will be made in this direction as long as the loss is concealed from view. An insurance system would disclose the extent and sources of this loss and make it a clear and definite expense to those who control industry and to the general body of society. The details of the system could be adjusted so as to make it to the advantage of employers to adopt every possible expedient for reducing the amount of unemployment in their enterprises.

Voluntary insurance. For many years the stronger trade unions in this country and abroad have had arrangements by which their members insure themselves by mutual action against the risk of unemployment. The system is a simple one: membership dues are set at a large enough figure to provide for the accumulation of funds from which to pay small pensions to members out of work. Since the benefits are distributed by the local unions, it is not difficult to safeguard the system against fraudulent claims for relief. As an additional safeguard maximum limits are set to the duration of relief in individual cases. Nevertheless this type of insurance is defective in many respects. The premiums paid by the members are not based on actuarial data and do not provide adequate reserves against times of widespread unemployment; in fact most of the systems collapse when the need for them is greatest, during periods of prolonged business depression. Furthermore trade union benefits are available only to a minority of the organized workers and not at all to the unorganized, so that the plan has slight effect upon the problem as a whole

The Ghent system. In many countries however these preëxisting relief systems of the unions have been utilized by government bodies as agencies through which to extend the principle of unemployment insurance at public expense. In 1901 the city of Ghent instituted what has come to be known as the *Ghent system* of unemployment relief, setting aside public funds for the payment of supplementary benefits proportioned to those of the trade unions. This system is administered through the machinery of the unions and is to a certain extent safeguarded against waste of public money by the requirement that no relief will be given except in

consideration of a proportional payment out of union funds. From Ghent the system spread to other municipalities of Belgium and through many other countries of Europe. By 1935, Belgium, Czechoslovakia, Denmark, Finland, France, Holland, Norway, Spain, Sweden, and Switzerland had adopted the practice of subsidizing these voluntary union plans with public funds. At this time about four million wage earners were covered by this type of insurance, paying from one-third to one-half of its cost out of their membership dues.

The Ghent plan has several serious weaknesses when viewed in relation to the whole problem of unemployment. In the first place, since the employers are not called upon to contribute to the relief of the unemployed, the system adds nothing to their inducement to keep their workmen employed. It does not of itself operate to reduce the unemployment evil. In the second place, it can be applied only to the organized workers. It leaves out of account the unorganized, who are frequently the most necessitous. Moreover in countries where the spread of unionism is opposed by the owners of industry the state cannot support the Ghent system of insurance without appearing to take part in this controversy, for to subsidize union unemployment benefits is in reality to assist the unions to enlarge their membership and to resist their employers.

Compulsory unemployment insurance. The weaknesses of voluntary insurance plans have led to the adoption of compulsory systems administered and partially financed by government. Compulsory unemployment insurance is by no means a novel experiment in social legislation. Throughout the world as a whole over 35,000,000 wage earners are now covered by this type of insurance. Eight European nations — Austria, Bulgaria, Germany, Great Britain, Ireland, Italy, Poland, and Russia — as well as certain governments in Australasia, have well-established compulsory systems. It is to the experience of these countries that we should look for definite information regarding the advantages and defects of the policy.

Compulsory insurance in Great Britain: Original plan. Great Britain was the first important country to adopt the policy of

compulsory unemployment insurance and the one country which has given the system thorough trial on a comprehensive scale and over a long period of time. The original law, passed in 1911, applied only to a few selected trades and covered about two and a half millions of the total fifteen million wage earners in the country. This law created a central unemployment fund, made up of equal contributions from the wage earners, the employers, and the public treasury. The wage earners' contributions were made weekly and varied in a declining scale according to the age and sex, but not the trade, locality, or wage rate, of the individual. These contributions were held back by the employer from the workers' pay and passed on to the government, along with an equal contribution from the funds of the business; these funds were again passed on to the Unemployment Fund, augmented by a grant from the treasury of the same amount as the workers' contribution. The payment of benefits was carefully safeguarded. No insured worker was eligible for benefit unless he had made at least thirty payments into the fund during the preceding two years. A simple device was adopted to provide proof on this point. Employers kept a book for each insured worker, in which he affixed a stamp each time a weekly payment was made. This book was given to the worker when he lost his job and constituted his proof of eligibility. The unemployed worker was required to register immediately in the nearest employment exchange and was obliged to accept any suitable employment offered him, the term "suitable" being carefully defined to protect the worker against being used as a strike-breaker or forced to accept conditions of employment materially below the standards of his craft. A waiting period of two weeks was required before any benefits were paid; and the duration of the benefit period was limited to one week for every six weekly contributions previously made by the individual and to a total of fifteen weeks in any year. The system operated to reduce unemployment, as well as to relieve distress, by giving the employer a refund of one-third of his payments for every worker continuously employed in each year.

The chief administrative features of the system have remained unchanged to the present time and are still in operation. The

rates both of contributions and of benefits have been altered from time to time. They have never been large; in 1937 the workers' contributions ranged from eighteen cents per week for adult males to nine cents for youths, and the weekly benefits from $4.25 for adult males and $3.75 for adult women to a bottom figure of $1.25 for youths, with small extra allowances for dependents. But these minor changes of detail have been of little consequence as compared with certain sweeping alterations in the scope of the system and in the rules of eligibility for benefits. For ten years the system operated successfully on an insurance basis; contributions into the fund had been large enough to meet all demands for benefits and leave a surplus of $1,000,000,000; benefits were not "doles" but businesslike payments to which the recipient was entitled by insurance contract.

Modified system after 1920. Beginning in 1920 and 1921 the situation underwent a radical change. Ten years later the fund was in debt to the government to the extent of $500,000,000; the stringent rules covering eligibility and limiting the benefit period had been greatly relaxed; the benefits paid out were made up in part of outright gifts or doles to the unemployed in violation of the strict principles of insurance.

Changes of policy with regard to two important features of the plan were responsible for the difficulties of the insurance system during the second decade of its life. In the first place, the number of workers covered by insurance had been greatly increased, without adequate provision of reserves against the increased risk. The amendment of 1916, which added some two million workers in war industries, caused little trouble, because there was no great increase of unemployment during the next few years. But in 1920 the system was abruptly broadened to include all wage earners receiving less than $1,250 a year except those in agriculture and domestic service, thus increasing the number insured to about twelve million people. This was done on the verge of a severe and prolonged business depression, so that there was no time to accumulate reserves for these new workers before great numbers of them were thrown into unemployment. Moreover as unemployment became more widespread, the standards govern-

ing benefits were relaxed through compassion or in response to political pressure. The waiting period was reduced from two weeks to six days; the duration of the benefit period was extended from fifteen to twenty-six weeks; the thirty contributions required for eligibility were not limited to the preceding two years but could be counted over any previous period. Even these changes did not suffice to meet the pressure on the system, and the labor exchanges were authorized by the government to go on paying benefits to unemployed individuals after the twenty-six week period had passed; that is, after the insured had exhausted his legal claim. Financial exigency led also to the abolition of the practice of refunds to businesses which provided stable employment, thus robbing the insurance system of its power to reduce the amount of unemployment. It must be said that the British government began this policy of paying excess benefits on the understanding that it was making temporary loans to the fund which would be repaid when the emergency was passed; but continued unemployment postponed the return to the original principles of the system until the government found itself in 1931 carrying an impossibly heavy financial load.

In that year drastic reforms were made. The benefits were reduced in amount and the contributions increased; eligibility rules were made more severe; and, most important, additional benefits to workers who had exhausted their legal claims were discontinued. These victims of prolonged unemployment were transferred to the charitable relief agencies of their local governments to be treated as cases of pauperism. The reforms had to be carried through in the face of violent political opposition, which has continued down to the present time. The insurance system is now reëstablished on a businesslike basis with all the "dole" features removed, but it has become an issue in partisan politics and may again succumb to popular pressure.

Lessons of the British experience. Examination of the British experience shows that compulsory insurance will work successfully under certain conditions. Many of the arguments against the policy ostensibly derived from British experience by opponents of unemployment insurance in America have been proved to be un-

founded. Even in the bad years following 1924, when the unemployed numbered continuously some two million people, there was only a trifling amount of malingering. The charge that insurance promotes shiftlessness among the working classes has been disposed of by the findings of several Royal Commissions which have investigated the problem. Investigations proved that 60 per cent of those unemployed at any one time had been idle less than one month, and only 25 per cent more than three months, while but a very small number had been out of work for six months or longer. There were cases of course of lazy or degenerate people, who welcomed the chance to live in idleness at public expense, but they were exceptional. The administrative efficiency of the system in normal times has been conclusively demonstrated, though it is important to note that this was due in large part to the existence of an efficient structure of labor exchanges. The experience shows too that, save in times of severe business depression, unemployment is an insurable risk susceptible of actuarial treatment and that insurance against it can be managed on sound business principles.

But the British experience also discloses certain dangers and limitations of unemployment insurance. The system will not withstand the strain of cyclical unemployment of a severe sort. There is danger at such times that the essential features of insurance will be abandoned in response to political clamor and that insurance will be transformed into a vast program for the relief of destitution out of public funds. Such relief is indeed an obligation of society, but it is a separate problem, to graft it on an insurance system will destroy the plan of insurance. Even if the political dangers could be avoided, it would still remain true that cyclical unemployment — the problem in its worst form — cannot be solved by insurance methods. It is doubtful whether adequate reserves can be accumulated to meet the drains of prolonged business depression. Even if enormous reserves are accumulated, they could not be turned into cash for the payment of benefits in times of depression without severe loss to the unemployment fund; and attempts to liquidate these reserves in an unsettled money market would tend toward further depression of general business conditions.

Unemployment insurance in the United States. Under the American Constitution the policy of unemployment insurance lies within the province of the several states. The federal government may take steps to encourage the adoption of state-wide systems by offering subsidies out of the national treasury; and if these subsidies are accepted, it may assume supervision of the standards and methods employed in the state plans. Until 1935, in the absence of such stimulation from the central government, the states had done very little; Wisconsin adopted an unemployment insurance law in 1932, and New York did likewise in 1934. At that time similar bills were pending in other state legislatures. This situation was abruptly altered by the passage of the federal social security act in August, 1935.

This law embraces a broad program of social legislation, including provisions for old age pensions, care of the blind and crippled, maternal care and child welfare, as well as compensation for the unemployed. Some of the new social policies were put under the administration of existing agencies, but to handle the new insurance and benefit programs the Social Security Board was created, consisting of three members with headquarters and a large staff in Washington and twelve regional offices in various parts of the country. We cannot attempt discussion here of all the features of the social security act but must be content to give attention only to its provisions for unemployment insurance. The law appropriates $4,000,000 during the first fiscal year and $49,000,000 in each succeeding year to be allocated to the states as a contribution toward the administrative cost of such insurance systems as they may set up. As an additional inducement to the states, the act imposes a tax on the payrolls of every business employing eight or more workers, at rates rising to 3 per cent in the third and each succeeding year, and provides that the employer may offset against this tax 90 per cent of any payments he makes to compulsory insurance funds under his state law. The law allows the states complete discretion in setting up systems of insurance suited to their own requirements or rejecting the policy outright; but the law requires that to qualify for grants from the federal fund the state system must meet certain tests with regard to its administrative and

benefit features, that funds accumulated by the states must be deposited in the Federal Unemployment Trust Fund of the United States Treasury to be invested in federal securities, and that the state system must report regularly to the Social Security Board and hold its records continuously open to the board's inspection. The act was declared constitutional by the Supreme Court in 1937.

At the present time the forty-eight states, the District of Columbia, Alaska, and Hawaii have adopted approved unemployment insurance laws. The different systems exhibit such great variety in matters of detail that it is impossible to do more than summarize their chief features. In all cases certain groups of workers are excluded by definition: agricultural labor, domestic servants, public employees, employees of religious, charitable, and educational institutions, and some others of minor importance. Moreover *in all but ten states workers in small concerns — usually defined as consisting of less than eight employees* — are not insured. It is estimated that some twenty million wage earners in all are covered by the insurance systems. In all but seven states, where the workers are required to make small contributions, the cost is borne by the employers in the form of payroll taxes. All the laws require a waiting period after the loss of employment before the benefits begin; all set a maximum period, frequently sixteen weeks, on the duration of benefits, requiring a minimum period of re-employment before additional benefits will be paid; and all adjust the benefits to the previous rate of earnings with a stated maximum, usually fifteen dollars a week.

These insurance systems have been in operation for too short a time to be tested by experience. In fact benefit payments were not available in any state until February, 1938, and it will not be until July, 1939, that all the systems become effective. No real appraisal of the policy is possible at the present time, although certain misgivings may be voiced regarding it. It is arguable for example that the policy of exempting the workers from any contribution, as is done in forty-four of the fifty-one systems, is unsound. Granted that the burden placed on the employers will be shifted at least in part to the workers, still it is probably not wholesome that the cost of insurance to the beneficiaries should be concealed

from sight. This feature of the American plans is no doubt attributable to the fact that the policy was pressed for adoption at a time when wage earners in general were receiving small and uncertain incomes, and millions were without any income at all The adoption of the policy at a time of deep business depression is unfortunate for this and other reasons. It launches the program at a time when it must bear excessive financial strain and when the difficulty of adhering strictly to sound insurance practices is the greatest. There is danger that the added burden on employers already struggling to avoid insolvency may retard business recovery and even cause increase of unemployment.

L

OTHER LABOR PROBLEMS

Hours of labor: The shorter work day. During the early stages of the factory system the typical work day embraced all the daylight hours. A work day fourteen hours long was not unusual. As a general rule no thought was given to the problem of hours save that the managers of industry attempted to keep the machines in operation as long as human endurance would permit. This condition lingered on in some branches of industry until very recent times, as for example in steel manufacture in America, where as late as 1922 there were blast furnaces in which as many as 69 per cent of the workers labored twelve hours a day seven days each week. But in industry as a whole a strong trend toward the shorter work day became apparent over a century ago and has continued down to the present. The shortening of the work day has been accompanied in recent years by a further reduction of hours of work by cutting down the number of work days in the normal week. The following table shows the trend of hours per day since 1840, and the shortening of hours per week since 1890.

NORMAL HOURS IN MANUFACTURING AND MECHANICAL ESTABLISHMENTS UNITED STATES [1]

Year	*Average normal hours per day*	*Average normal hours per week*
1840	11.4	
1850	11.5	
1860	11.0	
1870	10 5	
1880	10.3	58
1890	10.0	57
1900	9.8	55
1910	9.4	50
1920	8.5	49
1930	8 3	45

[1] Taken from C. R. Daugherty, *Labor Problems in American Industry*, Revised Edition, 1938, p. 192.

The depression following 1929 produced an abnormal situation as regards hours of labor, since many of the policies adopted by individual employers or governmental bodies to combat unemployment were of the share-the-work variety, involving part-time employment for many workers. In response to these influences there has been a strong movement toward a five-day week with not more than eight hours per day. In 1931 less than 800 concerns with a total labor force of some 700,000 were known to be operating on the five-day week schedule; in 1935, 250,000 firms employing 12,000,000 had adopted this schedule. Of course too much reliance must not be placed in these developments during depressions as forecasting the future trend.

The forces which have called into existence the pressure for the shorter day are not wholly economic in character. They spring in part from the improvement in general culture which has accompanied the material progress of the past century. Diffusion of education among the common people and the multiplication of their social and political activities create desires which can be satisfied only on condition that the average man enjoys a considerable period of freedom from the routine tasks of industrial employment. The services of libraries, museums, and amusement parks, provided free of charge by modern society, and the many forms of cheap commercial entertainment stimulate this desire for leisure.

The demand for the shorter day is also in part a product of the peculiar economic conditions of the modern factory system. The truth is that the productive life of the wage earner today is often a monotonous and joyless experience In such a situation there is no stimulus to effort except that afforded by the wage rate and no immaterial benefits in the form of satisfaction from work well done. The control of productive enterprise is lodged in the hands of those who own the capital instruments or their representatives; the workmen have no share in the planning and direction and can feel none of the joy of achievement which comes to the successful enterpriser. Finally fatigue increases progressively as the strain of effort is concentrated on a narrower range of muscular and mental faculties. Continuous labor of a certain duration, when it calls into play a

large variety of the worker's powers, is not so exhausting nor so repellent as labor of the same duration performed by a limited set of muscles. Those conditions grow more pronounced with the development of the factory system and the progress of machine technique. It seems probable that the productive life of the industrial worker must remain a dull and routine performance, forcing him to rely for self-expression upon ample hours of leisure spent outside the walls of the factory.

Relation of hours to productivity. The desirability of the shorter day as tested by considerations of general moral and social welfare is granted by most people. But every attempt to reduce the hours of labor raises the question of what wages can be paid for the shorter service. When the workmen press for a reduction of hours, they invariably demand that the hourly rate of wages be raised so that their earnings will remain unchanged. If we review the last seven decades, we find that a decline of about two and a half hours in the average work day has been accompanied by a rise of real wages for all groups of workers; but this cannot be attributed merely to the shortening of the work day, since other forces, particularly the improvement of machinery, were operating to increase labor's productivity.

Within limits a reduction of hours has been found to increase the productivity of labor. It has been proved by the experience of many different enterprises in Europe and America and established as a rule of wide applicability by researches carried on by the British Government and by various bureaus of investigation in America that a shortening of the more extreme work day, a fall of hours from twelve to eleven or ten, results in an actual increase in the total output of labor[1] To be sure, this effect of shortening the long work day is not uniform for all grades of workers. It is most noticeable in the case of those occupations which demand continuous mental alertness combined with manual dexterity and least so of the automatic types, such as the tending of machines whose pace is governed by the speed of machinery. But even the most

[1] The evidence in support of this conclusion is summarized in the *International Labor Review* for April and November, 1924 See also Vernon, *Industrial Fatigue and Efficiency*, London, 1921.

unskilled labor responds to a reduction of an extremely long day by some increase of output.

There are many reasons for this. When the length of the work day is excessive, the period of rest is not sufficient to counteract the effects of fatigue. In extreme cases accumulated fatigue predisposes the workman to sickness, which results in lost time, and where this consequence is avoided the workman tends to be slack in attendance through voluntary idleness. It has been found that the rate of accident varies directly with the length of the work day, reaching a maximum during the late hours, and also that it rises with the days of the week, thus disclosing another waste due to fatigue. Waste of materials, in the form of defective output, waste of power, and necessity of more labor of superintendence, through increased slackness of the workers, also result from protracted hours of labor. A saving of these wastes can be effected by shortening the hours of work to avoid the worst effects of fatigue. It is probable that the ten-hour day is more productive at the present time for all grades of labor than is the twelve-hour day and that for the more skilled groups an eight-hour day is at least as productive as ten hours of work.

But there are certainly limits beyond which reduction of hours cannot be carried without diminishing the total output of the workers and consequently causing a fall of real wages. It is difficult to determine these limits with precision, since they differ for each of the various grades of labor comprised in the working force of the typical industrial establishment. But to disregard them when reducing the hours of work is to diminish the material prosperity of the workers and of society as a whole. On this point we have evidence in the trend of real wages as compared with the length of the work day during the first two decades of the present century, when the shorter work day was gained by many of the workers at a sacrifice in terms of material income. And the conclusion is strengthened by investigations made of the effects of the shorter day on the productivity of certain grades of labor. It has been demonstrated that in many cases the gain in hourly productivity resulting from a substitution of the eight-hour for the ten-hour day was not sufficient to maintain the former level of total output.

Effects of technological advance. Each new invention and each improvement in organizational or managerial efficiency tends to increase the product of the average worker in a given length of time. Technical progress which is comprehensive, applying in general throughout the whole field of industry, must have the ultimate effect of expanding the material product of all groups of workers in a work day of given length. When real wages have thus risen high enough to establish a relatively comfortable standard of living for the higher groups of workers, each successive advance in industrial technique presents these workers with a choice between two ways in which to take their share in the general improvement. They may elect either to receive their former rate of real income and take their gains in shorter hours of work or, working the same number of hours as formerly, to raise their standard of material comfort.

This choice has been presented to the wage-earning population, without their knowledge, in the course of the rapid industrial evolution of the last three-quarters of a century. In many cases their actual choice has taken a middle ground between the two extremes. As their bargaining power has increased, the workers have pressed for a shorter work day and for a simultaneous increase of wages. That they have won both concessions is to be attributed in part to the positive increase of productivity resulting from reduction of long hours which we have discussed, in part also to the fact that these demands were made at a time when industrial progress was working in the direction of increasing the hourly productivity of labor. For some groups of workers the shortening of the work day has maintained their productivity at its maximum; for others it has resulted in an output below the maximum of which they are capable on the present stage of the industrial arts, though higher than was possible on earlier stages.

There is reason to believe that the rapid pace of technological progress which marked the years prior to 1929 will continue when normal conditions again prevail. If this should be true, the programs for a shorter work day now under discussion — programs which would reduce weekly hours to forty or less — may prove not impracticable. Such progressive reduction of hours will have the incidental advantage of retarding the spread of technological unem-

ployment. Their general social effect may be good in other respects, particularly if society devises more wholesome and beneficial uses for the increased leisure time of the people. These gains will be made however at some cost to the total material income of the wage-earning classes, a consequence of the shorter work day movement which its more enthusiastic advocates are prone to overlook.

Child and woman labor: Modern aspects. The problem of woman and child labor, like that of the length of the work day, can be called a modern problem only if we qualify our terms. The lot of the women and children of the poorer classes in society has always been one of toil, and it is probably true that both in respect of the conditions under which they work and of the rewards for their labor the modern world deals less severely with these more defenseless of its members than did society in its earlier stages. But it may be said nevertheless that the labor of women and children is a modern problem, not because it is more extensive or more degrading than formerly, but because its effects are different.

To grasp the significance of such a problem as this, at any stage of social development, it must be studied in its relation to the customs, institutions, and ideals which comprise the social structure of the time. It is because this relationship has been changed by the revolution of the social order which produced the modern industrial world that we are correct in speaking of the labor of women and children as a modern problem. The change has been due in part to an alteration in the ideals and principles which govern our political and social life. The tests of enduring welfare accepted in a democracy may be quite different from those which obtain in an aristocratic order. Conditions which deny to children opportunities for wholesome physical development and which interfere with the normal functions of motherhood or contribute to the spread of vice among unmarried women are recognized as constituting a problem of obvious social import. It is from the standpoint of its relation to the general welfare and to the ideals of the modern world that we must appraise the significance of child and woman labor.

But there is another way in which this problem differs from that of an earlier order; namely, in its effects upon the welfare of male workers. In earlier times the labor of women and children was

supplementary to, not in competition with, that of men. With the perfection of machine technique processes were so far simplified that jobs formerly monopolized by men could be filled by children or by women who possessed no considerable skill or physical strength. At once craft lines began to dissolve; cheap labor began to compete for employment with the skilled workers; women and children were in competition with men. This threatened a lowering of the standard of life and an undermining of the security of the male workers. With the progress of machine production this condition spread from one industry or process to another, until today it constitutes one of the major forces shaping the policy of the male workers.

Extent of child labor. American official statistics of child labor do not exist for the period prior to 1870. The United States census of that year reported 739,164 children between the ages of ten and fifteen at work. The numbers in succeeding decades are shown in the following table.

CHILDREN AGED TEN TO FIFTEEN GAINFULLY EMPLOYED
1880–1930 [1]

Year	*Total*	*Percentage of all gainfully employed people*
1880	1,118,356	6 4
1890	1,503,771	6.5
1900	1,750,178	6 0
1910	1,990,225	5 2
1920	1,060,858	2 6
1930	667,118	1 4

The figures given undoubtedly omit many children illegally employed and others not reported to the census-taker by their employers. Moreover if the ages sixteen and seventeen years were to be included in the count, the numbers would probably be more than doubled. Nevertheless the outstanding feature of the table is the rapid decline of child labor since 1910, both in the aggregate and relatively to the total working population. Agriculture has always been the chief employment of child labor In 1930 45 4 per cent of all gainfully employed children were at work in agriculture, as compared with 21.7 per cent in manufacture, 3.5 per cent in trans-

[1] Adapted from *Census of Occupations*, Abstract Summary, 1932, pp. 3–5, U. S. Bureau of the Census.

portation, 8.8 per cent in trade, 6.7 per cent in domestic service, and 8 per cent in clerical occupations.

The very recent trend of child labor is rather difficult to measure. Under the NRA codes the employment of children under sixteen was universally prohibited in industry, while some hazardous lines of occupation set the minimum age at eighteen years. There was an abrupt decline in the number of child wage earners estimated at two-thirds of the former total. A federal law of 1938, discussed below, has continued the restriction of child labor begun under these codes.

Effects of child labor. The evils of child labor do not require elaboration. The injurious moral effects which flow from an imperfect education, a denial of the period of play and recreation indispensable to mental and physical health, and contact with an environment which is often contaminating are apparent to all. There are as well certain economic effects which react unfavorably on the material welfare of society at large. The wage-earning child is restricted in his choice of occupation to lines in which there is no hope of considerable advancement in earning power. Even when he is not physically crippled by over-strain or by abbreviation of the normal period of rest, he is denied opportunity to acquire the skill demanded of those in the higher wage groups. At maturity he is usually caught in a "blind alley" occupation or recruited to the ranks of the already overcrowded group of unskilled manual labor. This involves a loss not only to the victim of child labor but to all the rest of society's members. For as we have seen in connection with our study of differences in wages, the total income of society would be increased if it were possible to transfer men from the relatively overcrowded groups whose marginal productivity is low to the less crowded occupations where marginal productivity is higher. Any device which hinders this transfer not only tends to perpetuate the inequalities of income as between the different classes in society but also prevents a possible net increase in the total income fund.

Child labor is one of those social problems which seem unable to work out their own correction. In some forms of occupation children can be used in place of men at a saving in cost to the employer.

Since they do not rely wholly upon their earnings for support, their wages are lower than the minimum at which it would be possible to maintain an adequate supply of self-supporting adult workers. In this there is an inducement, heightened by the pressure of competition, for the enterpriser to employ children wherever possible. This tendency is not counteracted by opposition from the parents. In very many cases, either through the pressure of their own poverty or through blindness or indifference to the welfare of the children, parents willingly consent to industrial employment for their children. The children themselves are obviously unable to resist. Nor do they show a disposition to do so. Quite naturally they rebel against the discipline of the school and, tempted by the prospect of a position of independence, are glad to engage in wage labor at the sacrifice of their future. When all parties to the contract coöperate thus willingly to increase the spread of child labor, the only remedy is that of restrictive legislation.

Legal regulation of child labor: In Great Britain. Early attempts to regulate child labor by law had to fight their way against the opposition of prevailing individualism. The first act of Great Britain, passed in 1802, applied only to pauper children, prohibiting their labor in factories if below the age of nine years, and if older prohibiting their employment for more than twelve hours a day or at night. The mere mention of these limits is enough to show the gravity of the problem of child labor during the early stages of the Industrial Revolution. Opposition even to mild limitation is shown in the series of amending and supplementary acts which were necessary before anything like thorough regulation of child labor could be accomplished. Not until 1833 were all textile mills brought under legal regulation. The law of that year prohibited the labor of children under the age of nine and provided that children between the ages of nine and thirteen should work only eight hours a day and those between thirteen and eighteen only twelve hours, with no employment at night. The factory acts were extended to all large industries in 1864 and to smaller concerns in 1867. In the meantime an act of 1842 had prohibited the labor of children under ten years of age in mines, which limit was later raised to thirteen years.

Without following in detail the development of this legislation, we may note the progress of thought on the subject by comparing with these early standards the terms of Great Britain's latest statute. By the law of 1920 fourteen years is made the minimum age for industrial employment for all children; the length of the work day for young persons between sixteen and eighteen years is limited to eight hours; no work is allowed at night for any wage earner below the age of eighteen. Other European nations of advanced industrial civilization have approximately the same standards. These laws are sanctioned by heavy penalties and administered by an efficient system of government inspection, which makes evasion extremely difficult.

In America. The regulation of child labor in the United States lies in the hands of the states, and practically every state has a law on the subject. Thirty-three states set a minimum of fourteen years, four a minimum of fifteen years, ten of sixteen years. Forty states set a limit of eight hours to the working day for children under sixteen. Only one state fails to prohibit night work for children, either entirely or in selected industrial occupations. The laws do not apply to work in agriculture or domestic service, and only a few of the states have attempted to regulate the street trades of children. The practical effect of these laws varies, not only because of their different standards but also because of the wide difference in the rigor with which they are administered in different states. In general the Southern states, where the factory system is of late development and where a necessitous population with low living standards creates a large supply of child workers, have the lower standards of regulation.

Lack of uniformity has been found to affect injuriously those units of a competitive industry which are located in the states having the more rigid standards of regulation, and this fact has tended to arouse opposition to child labor laws in these advanced states. Consequently various attempts have been made to bestow on the federal government power to impose uniform standards upon the country as a whole. A federal law passed in 1916 attempted to prohibit the labor of children under fourteen, to limit the work day of those from fourteen to sixteen, and to prohibit the labor of chil-

dren at night, by excluding the products of industries violating these standards from interstate commerce. When this law was declared unconstitutional in 1918, Congress passed another statute embodying the same standards and attempting to regulate child labor by applying prohibitive taxes to the concerns violating the standards. In 1922 this law also was judged unconstitutional. Another attempt to give the federal government authority over child labor took the form of an amendment to the Constitution adopted by Congress and submitted to the states for ratification in 1924. This amendment would bestow on Congress power to prohibit or regulate the labor of children under eighteen years of age; but ratification by the necessary majority of the states has not been obtained.

The latest attempt to prohibit child labor by federal statute was embodied in the fair labor standards act of 1938. With numerous exceptions, industries engaged in interstate commerce are prohibited by this act from employing children under sixteen years of age, while in hazardous industries the age limit is set at eighteen years. This act was approved by the Supreme Court in June, 1938, but it has not been in operation long enough to justify an opinion as to its probable effect on industry.

Labor of women: Extent and character in the United States. The rapid increase in the number of wage-earning women in the United States during recent years is shown by the following table.

GAINFULLY EMPLOYED WOMEN AGED TEN YEARS AND OVER
1880-1930 [1]

Year	*Number*	*Percentage of all women of age group*	*Percentage of all employed persons*
1880	2,647,000	14 7	15 2
1890	4,005,000	17 4	17 1
1900	5,319,000	18 8	18 3
1910	8,076,000	23.4	21 2
1920	8,550,000	21 1	20 5
1930	10,752,000	22.0	22.0

For the period as a whole the aggregate number of women wage earners increased fourfold, outstripping by far the rate of increase

[1] Taken from *Census of Occupations*, Abstract Summary, 1932, pp. 3-5, U. S Bureau of the Census.

of men workers. This increase continued at a rapid pace up to 1910, after which year the expansion of numbers, though continuous, was at a diminishing rate as shown by the percentages in the final column of the table. In 1930 over a fifth of all wage earners in this country were women.

An examination of the distribution of women workers among the various occupations shows that there are now few types of employment which women do not share with men. If the occupations calling for special muscular powers are left out of account, all other broad divisions of occupation will be found to employ women in large numbers. An analysis of the 1930 census of occupations gives the following information on this point.

PER CENT DISTRIBUTION BY SEX AMONG THE PRINCIPAL OCCUPATIONS

Occupation	*Male*	*Female*
All occupations	78 0	22 0
Agriculture	91.3	8.7
Manufacturing, total	86.6	13.4
Clothing factories	29.8	70.2
Textile mills	57.7	43 3
Trade	84.2	15.8
Professional service	53 1	46.9
Domestic service	35 8	64 2
Clerical	50.6	49.4

Despite this widespread distribution, competition for employment between women and men is an important factor in only a relatively small number of occupations. If the major groupings of occupation given in the table were broken down into a more refined classification, as in the case of clothing factories and textile mills, this concentration of competition would be clearly shown.

Causes of the increase in woman labor. To some extent the greater employment of women is due to the progress of machine technique, to which reference has been made at other points in our study. Quickness and deftness of touch take the place of muscular strength and endurance as qualifications in the labor supply of those industries which are equipped with automatic power-driven machinery. This makes it possible to substitute the labor of women or children for that of men.

But the movement is not wholly accounted for by reference to

this factor. Mechanical equipment is of chief importance only in manufacturing, where, as we have seen, the labor of women has not risen to such formidable proportions as in other lines of occupation. Moreover the possibility of adapting the job to the characteristics of the woman worker does not of itself suffice to explain why the women of the modern world are more willing than in former years to seek employment as wage earners. For an explanation of this fact we must look to tendencies of broad social significance in the customs and habits of the people. The postponement of marriage, the removal of customary restraints which formerly circumscribed the status of women in the home and the neighborhood, and the increase in amount and change in character of the education offered to women are all forces working toward an increase in the disposition of women to seek positions of economic independence. Founded as it is on such fundamental social forces, this tendency toward increased employment of women is probably irresistible.

There are unquestionable gains resulting from the influx of women into industrial occupations. For the women themselves, this change in the conditions of their lives has meant escape from the inanity and futility of woman's existence in the Victorian age and an opportunity for self-expression and self-development which must contribute to a richer and more significant life. Society has gained materially, through the substitution of productive for meaningless tasks for its unmarried women, and spiritually, through the release of energies formerly thwarted by the bonds of custom and tradition. But these benefits have not been unmixed with other consequences. There are obvious moral and health hazards which society cannot lightly disregard; there are influences upon the home — still the basic unit of our social structure — whose general social consequences are not wholly desirable. We shall pass over these aspects of the matter and consider only certain problems of a strictly economic nature.

The problem of women's wages. Woman labor presents two interconnected problems with respect to wages: the amount and adequacy of the wage received by women workers, and the effect of their competition upon wages in general. All investigations agree in the conclusion that the pay given to women in industry is very

low. In 1914 the average weekly wage of women in factories was $7.75; in 1921, $15.63; in 1929, $17.61; in 1937, $16.95.

No argument is needed to prove that such wages as these are inadequate to provide a wholesome standard of living. It is true that many of these workers did not depend for support solely on their wages; many were married or belonged in other capacities to a household supported in large part by the earnings of a husband or father or by the joint earnings of different members of the family. This aspect of the problem however is not so pertinent to our inquiry as is the comparative level of wages for women and for men.

Comparative wages of men and women. All the studies of this question concur in the finding that women are paid less than men for the same task and in the same industries One investigator comparing the levels of wages in 1899 and again in 1923 discovered that in both years women's annual earnings were 54 per cent of men's in comparable employments. The National Industrial Conference Board found this same ratio in 1929 between the wages of women and men in skilled employments; while in unskilled occupations women received in that year 72 per cent as much as men. The best available information indicates that the differential between men's and women's wages in 1937 was about the same as in 1929. Some part of this discrepancy is to be explained by the inferior productivity of the woman worker; but when all possible allowance has been made for this factor, it still remains true that women are consistently underpaid as compared with men for no other apparent reason than that they are women.

This systematic underpayment of women cannot be accounted for except on the ground of their weaker bargaining power. As compared with men, women are notably weak in asserting their claims against the employer. We have seen that they are concentrated for the most part in certain lines of occupation. Here they meet severe competition from a fluctuating group of women laborers who seek part time or temporary employment as a supplement to their labor at home. Again they are even more immobile than the men; it is difficult for them to play off one employer against another when this involves a change of residence. In many

cases also they are not self-supporting and make no pretense of demanding a "living wage"; the minimum wage which they will accept is lower than that of men, because they are partially subsidized by the male wage earners of their families. Those young women who look upon their employment as a temporary expedient to fill in the time before their marriage are not always intent upon exacting the highest possible price for their labor. Finally they have been until lately unorganized and are still to some extent unorganizable; their bargains are made as individuals, and they labor under all the handicaps which the isolated worker must bear in the modern labor market.

The "equal pay" doctrine. Men workers look upon the competition of underpaid women as destructive of their own standards of wages, as indeed it is. In those industries where the two groups meet in contest for employment, the situation is very similar to that encountered by the self-supporting producer in a market filled with subsidized goods which compete with his own. Such competition, whether in the sale of goods or of services, makes for low prices for all sellers; and in the labor market, if the proportion of women workers is large enough, their low wages will determine the conditions on which the male workers can find employment. In their struggle to protect themselves against this situation, the men workers have devised the slogan so frequently found in trade union literature: "Equal pay for equal work."

Two very different meanings may be attached to this demand that the wage rate for women be the same as that for men. In the first place, it may mean that the payment for a given product should not be less for women, solely because they are women, than it is for men. Since women are admittedly less productive in most occupations than men, the equal pay slogan in this sense would result in lower day wages for women. The second meaning resolves the slogan into a demand that the daily or weekly wage of women be the same as that of men, regardless of the relative productivity of the two groups of workers. If applied in this sense the result would be higher wages in relation to productivity for women than for men. As advanced by most of the labor unions, the demand for equal pay has the second of these meanings. Its intention is, not to raise the

wages of women, but to exclude them from employment. If employers are required to pay the same wage to two groups of different productivity, it is obvious that they will not be inclined to employ the less efficient workers at all. On economic grounds there is little to be said in defense of the doctrine of equal wages when applied in this sense.

Leaders among the women in industry are prone to adopt the same impracticable definition of the equal pay principle. Obviously it is to the disadvantage of the women workers to demand conditions of employment which give the employer a preference for the labor of men. On the other hand, it is both just and practicable to pay women at the same rate in proportion to their relative productivity as men, and devices which increase the bargaining power of women will tend to produce this result. Past experience has shown that unionization for collective bargaining is effective in such a situation.

Legal regulation of woman labor in America. Aside from the minimum wage laws of certain states, which were discussed in another place,[1] laws bearing on the labor of women have been devoted primarily to setting a maximum limit on their hours and to prohibiting or limiting woman labor in certain unwholesome employments. At the present time only five states fail to set some sort of legal limit on the work day for women. Fifteen of the states set the maximum at forty-eight hours a week; in eleven states the maximum is fifty-four hours; a few allow a work week as long as sixty hours. The constitutionality of these laws, which for many years was in question, was decided in 1908 by the action of the United States Supreme Court in upholding the ten-hour law of Oregon. In general women are excluded from mines and other unusually hazardous occupations. Twenty states forbid the employment of women at night; and many have special laws safeguarding the health of the woman worker. Little is done to regulate night work. The laws specifically exempt farm work and domestic work from regulation.

Old age dependency: The problem. Poverty among aged people is a labor problem in that it is to be found principally within

[1] Chapter XIX.

the wage-earning class and symbolizes the inadequacy of the average wage income to provide lifetime support for the worker. The amount of dependency in old age also reflects other aspects of the modern industrial era: the relatively early age at which many workers are thrown into the discard by the production methods of the factory system, the incidence of accident and occupational disease, which renders many workers non-productive, the impoverishing effects of unemployment, and the like. The extent to which the inhabitants of the United States in the higher age groups must rely for support on charity or on the earnings of others is not fully appreciated by many people. A recent government survey disclosed the following situation in 1937.

PERSONS 65 YEARS OF AGE AND OVER[1]

	Number	*Per cent*
Total	7,816,000	100
Self dependent	2,746,000	35 1
On earnings 12.8%		
On savings 15 0%		
Dependent	5,070,000	64 9
On friends and relatives 44.6%		
On public or private agencies 20 3%		

In brief, out of every hundred people in this age group, sixty-five are kept alive by private charity or public relief.

Public provision for the aged: Variety of practice. Society has for long been making provision for the care of the indigent aged. The original system, familiar in this country, was to treat them as paupers and grant them poor relief through the agencies of local government, usually the county. The almshouse or "county farm," found in all sections of the United States, typifies this attitude of government toward its superannuated citizens. The practice is now in general disrepute, both because of its degrading influence on the aged poor and because it has been found to be extremely expensive in proportion to the numbers of people benefited by it.

Since the turn of the century most nations have taken steps to place relief for the aged on a more definite basis. The plans now in operation are of two distinct types, the *pension* and the *insurance* system. The former provides grants for aged dependents in the

[1] Bulletin of the Social Security Board, March, 1938.

form of weekly or monthly doles payable currently from the public treasury without previous accumulation of reserves. In other words, it is a form of state charity differing from earlier hit-or-miss systems of relief in that payments are made in the form of regular monthly incomes, as a matter of legal right to all aged persons, regardless of social class, on proof of need. An insurance system, on the other hand, is organized on the annuity principle, whereby the future incomes of the beneficiaries are accumulated in advance out of payments made into a reserve fund. If established on a sound basis old age insurance collects over the productive lifetime of the beneficiary contributions sufficient to cover the cost of the annuity. In some countries both systems are in operation at the same time, pensions being provided for those who are not covered by an existing insurance scheme.

Old age pensions. At the present time old age pension systems are in operation in about a dozen countries, including France, Great Britain, and the United States. In the latter country the policy is in the hands of the several states. Definite provision for old age pensions by state law began in the twenties; by 1935 twenty-eight states had such laws in operation, paying pensions totalling upwards of thirty million dollars to nearly a quarter million people. The plans all contained about the same features: the eligible age was sixty-five years, American citizenship and five years' residence in the state were required, the amounts paid were usually limited to thirty dollars a month, payments were made only on proof of need. This type of provision for the aged poor, developing slowly throughout the country, was given an immense impetus by the federal social security act of 1935.

The social security act provides for old age insurance as well as for assistance to the pension plans of the states. The latter feature is covered by a section of the act empowering Congress to appropriate $49,750,000 in 1935–36 and in following years as much as may be needed to make grants to states which set up old age pension laws approved by the Social Security Board. The approval feature has been used to strengthen the administrative machinery of the state laws and to provide complete coverage of all indigent aged people within the state who qualify as citizens and residents.

When approval of the state law has been granted, the federal treasury undertakes to bear half the expense of the pension system up to a maximum of fifteen dollars a month per person. At the present time all but one of the States, as well as Alaska, Hawaii, and the District of Columbia, have approved pension schemes in operation. These schemes embody the features of the earlier state plans summarized in the preceding paragraph. In 1938 a million and a half people were receiving pensions totalling upwards to $270,000,000 annually.

Old age insurance. As distinct from old age pensions, old age insurance had been set up in many other countries before it was adopted in America. All told, thirty-two nations, including France, Germany, and Great Britain, were providing such insurance in 1938, either to all aged people compulsorily, as in Germany, or to those who voluntarily entered the scheme, as in Japan. In the United States until 1937 wage earners in private employment had available in old age only the pension plans described above, although compulsory contributory insurance had been established for the public employees in many cities and states and in certain services of the federal government.

United States social security act of 1937: Description. The old age insurance provisions of the social security act, in effect in 1937, provide a nation-wide plan for the wage earners, with exception of certain categories of workers excluded by definition, under the direct administration of the federal government. In addition to governmental employees of all types, the following classes are excluded: agricultural labor, domestic servants, irregularly employed casual workers, the employees of religious, charitable, and educational institutions, and marine and railway workers. All other hired workers, regardless of the amount of wage or salary — the general manager as well as the manual laborer — are embraced in the system.

The details of the plan are complex, since they provide for a transition period during which the contributions to the reserves gradually increase until the scheme goes into full operation in 1949; but the essential features are these: contributions are made in equal amounts by employer and worker in the form of a per-

centage tax on the payroll and the income. In 1949 the employer will be taxed 3 per cent of the wages of every insured employee, and the employee will be taxed an equal amount on his wage income; in the intervening years the taxes rise by graded amounts toward this percentage, beginning with one per cent in 1937. In practice the employer pays both halves of the contribution, deducting the worker's share from his pay envelope or salary check. The government makes no contribution toward the scheme. Funds raised in this way are carried in the old age reserve account of the United States Treasury to be invested in three per cent federal securities. Annuities are payable at the age of sixty-five, regardless of the financial circumstances of the individual, on condition that he ceases to draw wages at that time; as long as he continues in employment, the payment of the annuity is postponed. The amount of the annuity is to be a percentage of the total wages received by the beneficiary between 1937 and his attaining the age of sixty-five, with both minimum and maximum limits fixed so that no one receives less than ten dollars nor more than eighty-five dollars per month. Provision is made for a lump sum payment of three and a half per cent of the total wages in case of death prior to the age of sixty-five and for a return to the individual's estate of this amount less any benefits paid him when he dies after the age of sixty-five. Payment of benefits will begin in January, 1942, though of course the amounts will be small at that time because of the brief period during which reserves have accumulated. To give effect to this plan, the government has created a vast administrative machinery, in order to register the millions of workers involved and to set up a separate account recording for each the contributions received and payments made on his behalf.

The "reserve" feature. This great experiment in social insurance raises many questions of a practical nature which we shall not consider here. There is however one problem of basic importance to which attention should be given. The essence of the plan, considered by its advocates to be of outstanding merit, is the apparently businesslike principle which it embodies. The whole scheme is described as self-financing in that collections are to be made from each beneficiary equal, together with compound interest, to the

benefits ultimately paid him. Since each beneficiary must pay premiums during his productive years without drawing benefits, the whole scheme involves the creation of enormous reserves in its initial period. It is calculated that the system will have attained maturity about the year 1980. At that time and for the indefinite future annual outpayments in benefits will be balanced by annual receipts from taxes on employers and workers plus the income from the accumulated reserves, the total reserve being estimated for that year at forty-nine billion dollars. A moment's thought will make it clear that the financial integrity of the plan (in the business sense) depends on its success in keeping this tremendous reserve invested so as to return the calculated rate of interest -- three per cent.

It is noteworthy that the old age insurance reserve is to be invested in government securities; in other words, it will create a permanent national debt of the amount indicated. There are serious problems of public finance associated with such a forced growth in the government's debt; but for our purpose the important point is the misleading appearance of self-financing given to the scheme. Unless the government invests these large funds in productive enterprise, which is not only impracticable but undesirable, it will have only one source upon which to draw for the interest payments on the securities held in the old age reserve account; namely, the current tax receipts of the federal treasury. The outcome would not have been different essentially if the taxpayers had been called upon frankly to make up the difference between the income from the payroll taxes and the benefit payments guaranteed the beneficiaries. This effect of the plan will not become operative during its early years, when the contributions exceed the benefit payments and the reserves are accumulating, but only when income from the reserves must be drawn upon. It is estimated that eventually some one and one-half billion dollars annually will be required from the public treasury to balance the account.

These statements are not intended as a criticism of old age insurance as such but as an explanation of the real effect of certain administrative features of the plan adopted in this country. Old age insurance is too firmly established in modern countries and too

strongly supported by economic and humanitarian considerations to be open to serious challenge. The principle of requiring contributions from the beneficiary is sound. To require the employers to pay a share is probably defensible as an administrative measure, but it is somewhat misleading as to the final incidence of the cost. The employer will shift the tax, since it must operate as one of the regular costs of the business. This part of the contribution will eventually come to rest in large measure on the wage earners.

Industrial accidents: The problem. One of the principal costs involved in the modern system of industry is due to the temporary or permanent destruction of the productive power of the workers through injuries resulting from the normal hazards of their occupations. Only within recent years has serious attempt been made to estimate the magnitude of this cost. The official statistics of work accidents in the United States are even now so unreliable and incomplete that any statement of the total loss can be no more than an estimate. In some states no records at all are kept. Certain industries have accurate records of the fatalities resulting from the hazards of the occupation, though the less serious accidents which result in temporary disability are allowed to go unrecorded. Even on a conservative estimate the loss can be shown to be surprisingly large. To give the result of only one study out of many, the National Safety Council found that in 1929 industrial accidents caused 20,000 deaths. The number of minor accidents has been variously estimated, usually in the neighborhood of 2,000,000 a year.

In recent years strenuous efforts have been made to reduce this waste. Factory acts and compensation laws have made it obligatory or profitable for individual owners of business enterprise to safeguard the lives and health of their employees. Groups of employers in the major lines of industry have coöperated through their trade associations to perfect devices for the prevention of accident and to urge their adoption upon the more indifferent individuals among their number. In the United States this effort has given rise to a widespread "safety first" movement, which enlists the services of experts and adopts many ingenious devices to make its program effective. Similar movements have been under way for

decades in other countries, reaching their point of highest efficiency perhaps in Germany, which leads the world in accident prevention. There can be no doubt that these concerted efforts have served to prevent a vast number of accidents.

The unavoidable loss of accident. But the most impressive result of such efforts has been to demonstrate that this cost is on the whole unavoidable, an inevitable feature of our industrial system. After every possible precaution has been taken to prevent them, work accidents show a tendency to increase, both in total number and in relation to the laboring population. This unhappy result flows from the maladjustment of the human mechanism to mechanized industry. Modern machinery employs forces of great power and applies these forces to production with remarkable precision of tempo and rhythm. The brain and body of man have not developed controls of sufficient accuracy and tirelessness to enable the workman to adjust his movements perfectly to those of the machine. Clumsiness, inattention, and fatigue prevent such coördination of the human to the mechanical factor in industry as would make of the two a perfectly functioning single mechanism. Accidents occur when this maladjustment becomes too pronounced. The spread of machine technique therefore involves an increasing toll of accidents despite all efforts to prevent them.

It is obvious that all work accidents result in net social loss. Even disregarding the suffering of the victims and their families, society cannot but be injured by this continuous destruction and crippling of the units of an essential productive instrument. When a laborer is thrown into idleness, society will be deprived irrevocably of that contribution to its income fund which he would have made. Since accidents cannot be wholly avoided, the only effective remedy lies along two lines of development: reduction of the number of work accidents by applying the resources of science to their prevention and, when accident does occur, rehabilitation of the workman as speedily and completely as possible so that his productive energy may be placed once more at the service of society.

Distribution of the loss: Conditions of the problem. Though social loss from work accidents is inevitable, arrangements can be

made which affect its distribution. Such arrangements do more than raise the question of justice. It has been found that different methods of apportioning the loss from work accidents have different effects upon the number of accidents and upon the magnitude of the social loss involved in them.

The burden must fall in the first instance upon the injured workman and his family. Where the rate of accident differs strikingly as between different occupations, this hazard will be one of the factors in the distribution of the labor supply. If the workers were entirely free to choose their lines of occupation, and if they made this choice on the basis of cool calculation of costs and income, the more hazardous forms of employment would be relatively undermanned, and the wage rate would be relatively higher than in occupations which were less hazardous. In this way compensation for the risk would be worked out by the forces of demand and supply, and the prices of the different kinds of goods would carry the cost on to the consumer. It is probable that a partial adjustment was made in this way before society undertook to distribute the risk by collective action. But we have seen that the distribution of the labor supply is not controlled principally by rational choice but rather by conditions of birth and tradition. The supply of labor in the hazardous occupations, as compared with other lines of work, is not materially limited by the risk of accident, and the wage rate, though it may contain a partial compensation for this factor of cost, does not suffice to lift the burden entirely from the workman.

Damages under the common law. Before the passage of the compensation acts, the injured workman in the United States was entitled to sue the employer for the recovery of his loss. In the courts the justice of his claim was tested by the principles of the common law defining the liability of employers. These common law principles were evolved at an earlier stage of industrial development to govern the simpler relations of master and servant, and their maladjustment to the conditions of the modern world placed serious obstacles in the way of recovery by the workmen. Stated briefly, the court attempted to discover who was to blame for the accident; if it could be shown that the fault was wholly the employer's, the workman was awarded damages; but if it appeared that the work-

man himself or someone else shared the blame with the employer, no matter in how slight a degree, the case was dismissed without compensation. This was the essence of the common law *doctrine of negligence,* which in all Anglo-Saxon countries determined the employer's liability to indemnify his workmen in case of accident in the absence of legislation to the contrary.

The difficulty under which the victim of accident labored in his attempt to shift the burden of loss will be apparent when we consider the defenses to which the employer could resort in answer to the workman's suit. In the first place, he could assert that the accident was due at least in part to *contributory negligence* on the part of the victim. The inability of the human organism to adapt its responses perfectly to the rhythm of the machine, coupled with the complicating effects of fatigue which we have already noted make it probable that most accidents are due at least in part to some mistake of the victim.

Failing proof of contributory negligence, it might be shown that the accident had been due to some fault of a fellow employee. The increasing interdependence of the personnel of the modern factory and the responsibility borne by those workmen who control powerful machinery which acts as a constant menace to the lives and limbs of their associates makes the *negligence of the fellow servant* an important factor among the causes of accident. If any element of blame could be attributed to a fellow employee, the employer was released from liability.

If these defenses failed, the employer might yet show that the accident had been due to a normal hazard of the business, of which the workman was aware when accepting employment. Under the doctrine of *assumption of risk*, the victim of accident had no claim against the employer for injuries resulting from these normal and known hazards. Considering how little choice the unemployed laborer has in the matter of accepting the job which offers, it is clear that this defense would in many cases be effective. It is not surprising that the workman was rarely successful in recovering indemnity for his injury.

Compensation laws. Modern compensation laws require the employer to indemnify the injured workman for his loss in wages

and medical care and to compensate the victim's family in case of fatal accident. These laws are based on a line of reasoning quite different from that which gave support to the common law doctrine of negligence. They give up all attempt to assess the blame for accident, recognizing that in the modern system of industry destruction of life and limb are inevitable normal costs. The financial burden involved in compensation for accident and death is accordingly placed solely on the employer. It is expected that this cost will become part of the enterpriser's expenses and enter into the price of the product, thus finding its way to the consumer, upon whom in the last analysis the burden should fall.

One of the greatest virtues of the policy of holding the employer financially responsible for accident is that it operates to reduce the number of accidents. By making the hazards of occupation expensive to those who control industry, the compensation laws create an incentive in the group which alone has the power to prevent them. The "safety first" movement and similar campaigns for the reduction of accident sprang up after society had determined to hold the employer responsible for the cost of injury received in the course of employment.

History of compensation laws and insurance. Germany led the way in legislation requiring the compensation of injured workmen. A law of 1884 gave effect to the principle that all employers should be liable to their workmen for partial compensation for wages lost through injury (two thirds of the former wage is paid to those permanently disabled) and to the dependents of the workmen for the loss of earnings resulting from fatal accident. Industrial accident is an insurable risk. The German system worked out a basis of insurance within each major group of industries and compelled the employers in each of these groups to form mutual insurance companies, to which the different business concerns paid premiums graded according to their accident rates. The whole system is supervised by public officials. Its efficient management soon led to a declining accident rate, which has placed Germany in the forefront of modern nations as regards this waste of industry. In 1897 Great Britain passed a comprehensive act requiring the compensation of injured workmen. This act makes no attempt at compulsory

insurance, merely placing on the employer the liability of paying one-half the former wage in case of non-fatal accident and a lump sum roughly equivalent to three years' earnings in the case of death. As a practical matter British employers almost invariably insure their liability in some private insurance company, though not required to do so. France and Belgium have similar systems. At the present time every industrial nation has workmen's compensation laws.

The first laws on this subject in the United States were declared unconstitutional, and it was not until 1909 that New Jersey succeeded in setting up a system of compensation which was acceptable to the courts. With the single exception of Mississippi, all the states and territories had passed laws on the subject by 1938, and the federal government had adopted compensation acts covering its own employees and the longshoremen and industrial workers in the District of Columbia. Twenty-seven of the laws, including that of the federal government, cover certain occupational diseases as well as accident.

The essential feature of these laws is their requirement of compensation by the employer to the injured workman or in cases of fatality to his heirs. Aside from this common feature the laws vary greatly in detail. In only twenty states does the compensation act require the employer to join any insurance system or become party to an arrangement for appraising amounts of compensation. In the other jurisdictions the employer is allowed, if he chooses, to fall back on the old common law but without the three defenses described above. Rates of compensation vary but are always a percentage of the injured worker's wages, paid in the form of life pension in case of permanent injury, as a capital sum or an annuity to the heirs in case of death, and as a single payment in case of temporary disability. Many of the laws provide remuneration for hospital and medical costs as well. The laws requiring insurance also vary in that some provide a state insurance fund requiring all employers to subscribe to it; others provide such a fund but leave the employer free to insure in private companies; and still others make no special provision for handling the insurance problem.

The system is by no means complete, even in the states which have adopted it. Certain categories of workers are excluded; farm laborers and domestic servants almost invariably, in some cases workers in non-hazardous industries, and all seamen, railway workers, and bus drivers in interstate transportation. In many states the workers in very small enterprises are not covered. Altogether it is estimated that about twelve per cent of the wage earners are without this protection.

LI

THE ORGANIZED LABOR MOVEMENT

Conditioning factors. Unions of workers for collective action are to be found throughout the world wherever the modern system of industry has taken root. Appearing first in England, the birthplace of the factory system, these organizations have spread to all industrial countries and have flourished in rough correlation to the extent of industrialization. Today we may mark their beginnings in countries — such as India and Egypt — which are only just now taking over the productive methods of the more advanced nations. According to the most reliable information available, the labor movement throughout the world as a whole numbered 44,000,000 members in 1929. Over 35,000,000 of these were in Europe; 7,000,000 in the industrial countries of America; 1,000,000 in Australasia; three-quarters of a million in Asia; and less than 100,000 in Africa. Not only is there this close correspondence between the extent of union membership and the type of industrial organization dominant in a country, but within the most highly industrialized countries, where the movement is strong, unionism *flourishes only among those classes of workers who are directly affected by modern methods of production.* In countries like England and the United States there exist great numbers of wage earners on the farms, in retail trade, in domestic service, and in other positions of semi-independence, among whom labor unions have been unable to gain a permanent foothold.

These facts indicate that the labor movement is an attempt on the part of great numbers of men to adjust themselves to certain conditions of life created by the modern system of industry. *These conditions have been suggested by our preceding study* and need only be summarized here. In the first place, this system tends *to separate the worker from ownership in the capital instruments* and the land with which he works. This destroys the position of the laborer as an independent enterpriser and reduces the mass

of men in society to something like a uniform economic status, welding them gradually into a distinct social class. Secondly it concentrates the workers in close contact with each other upon a limited geographical area. The discipline of the factory or mine, by accustoming the wage earners to a system of regimentation in which conscious coöperation is carried out under orders of management, provides conditions which conduce to voluntary group conduct. In the third place, the modern industrial system has released certain injurious economic forces which affect the workers in common and against which they are as individuals powerless to protect themselves. These forces have been sketched in preceding chapters; namely, unemployment, low wages, monotonous and hazardous forms of work, the labor of women and children — all problems whose solution is vital to the security and contentment of the mass of wage earners. The existence of these common problems supplies the motive for collective action among workers, as the regimentation of their productive life creates the condition which facilitates such action.

The labor movement in Great Britain: Origin. While space does not permit a detailed account of the rise of the modern labor movement, some reference to its history is required to make clear its essential characteristics. Unions of workers, organized with the intent of controlling wages and conditions of employment, began to attract public attention in England before the opening of the nineteenth century. These were at first isolated trade or craft groups — workers engaged in the same skilled occupation — whose predominant motive was to secure the enforcement of the customary and legal guarantees as to wages and the right to work which had come down from the medieval period of England's industrial life. This attempt to perpetuate the conditions of an earlier economic system failed; economic forces generated by the factory system swept away the monopolistic rights of the skilled craftsmen and rendered obsolete the apprenticeship rules upon which these rights were based. The old laws fixing wages and regulating employment were repealed, and a public policy was established which accepted as its basis the principles of individualism and competition which were becoming dominant in the eco-

nomic life of the country. In harmony with this newer viewpoint, all unions of workers for collective economic action were declared illegal by laws — the combination acts — passed in 1799 and 1800. The following quarter century witnessed a rigorous repression of those illegal combinations of workers which persisted despite the severity of the law. During this period the factory system was making rapid headway in England. Competition among the wage workers became increasingly bitter as the progress of machine technique undermined the security of the skilled trades and made possible the employment of women and children in processes formerly monopolized by men. Wages fell, and hours of labor and conditions of sanitation and safety in factories, mines, and tenements became intolerable. In self-protection the male workers in the different industries united to resist these destructive tendencies, directing their efforts toward an increase of their bargaining power in dealings with their employers rather than toward a reëstablishment of medieval conditions. Fines, imprisonment, expulsion from the country, punishments visited by the law upon the leaders of these unions, did not avail to crush the movement. When the combination acts of 1799-1800 were repealed in 1824, the law merely took cognizance of the fact that organization among the workers was a natural feature of the new type of industrial society.

Conflict with the law. From 1825 to the present day the formation of unions has been legal in England. However during the half century immediately following the repeal of the combination acts, the activities of each union were narrowly circumscribed by common law doctrines, which condemned as criminal conspiracies all agreements to raise wages and prices or to restrain trade. Since the unions were formed for the express purpose of winning better terms of employment, and that by collective agreement and collective action, their members were continually falling foul of this doctrine of conspiracy. Nevertheless the union movement spread with great rapidity. During the thirties it was caught up in the widespread social agitation of that decade and wasted its energies on support of schemes of social and political revolution. The collapse of the Chartist movement had a sobering effect upon

labor leaders. The labor movement turned its attention to the problems of perfecting its organization and improving economic conditions within the different trades. Isolated unions embracing similar grades of labor in different sections of the country began to coalesce into nation-wide organizations. As these unions grew in strength their normal policies of collective bargaining and their weapons of warfare, the strike and the picket, brought them more and more into conflict with the courts, which were prone to view all such group activities as evidence of conspiracy to harass the employer and to interrupt the free play of competitive forces in the labor market. Largely in response to this common problem the various unions of the country met in 1868 in a trade union congress to establish an alliance on a national scale. This first congress was attended by representatives of 118,000 unionists; in the congress of the following year 250,000 unionists were represented; in 1872 the number had grown to 375,000. Since 1872 the annual trade union congress has been an established feature of the labor movement of Great Britain.

Recent developments. The period from 1870 to the present day has been marked by several lines of development in the British labor movement. In the first place, there has been a steady growth in union membership and a continuous improvement of union structure in the direction of integrating the various groups into strong national organizations. By 1920 the labor movement had attained a membership of approximately eight and a quarter millions. The importance of this movement as a factor in British national life is indicated by the fact that it embraced in this year sixty-five per cent of all wage earners engaged in industrial occupations and over eighty per cent of the male workers in these occupations. During the prolonged depression following 1920 the British labor movement declined both in numbers and in bargaining power, but they have regained their strength in recent years with the return of prosperity. In the most important branches of industry; *i.e.*, mining, transportation, the mechanical trades, and the building trades, the workers are united in great nation-wide organizations, each of which includes virtually all the available labor in its particular department of industrial activity.

In the second place, the law has been remodelled along lines more and more favorable to trade union policy. A decisive step was taken by the government in 1875, when the conspiracy doctrine was modified to exempt the unions from all penalties for peaceful collective action. The unions, as corporate organizations, were no longer punishable for forming agreements to raise wages by means of their power to control the supply of labor nor for engaging in strikes in furtherance of this aim. When in the early years of the twentieth century the courts showed a disposition to weaken these legal guarantees by interpretations of the law adverse to union activity, the matter was set finally at rest by the trade disputes act of 1906. By the terms of this act the courts were debarred from entertaining suits against trade unions, as such, for any cause whatsoever. The individual members, like all other citizens, were punishable under the criminal law for violence to persons and property, but in their collective capacity they were immune from legal interference. The trade union act of 1927 somewhat weakened these immunities of the unions by prohibiting the general strike and providing for legal procedure against labor organizations for injuries resulting therefrom.

In the third place, the British labor movement has become a powerful factor in the political life of the nation. Following a fitful and sporadic attempt to found an independent labor party in the nineties, organized labor entered the political arena during the first decade of the twentieth century in a determined manner. During the next quarter century the Labor Party grew in power with greater rapidity than any other party in the political history of modern democratic nations. In a three-sided election early in 1924 the Labor Party, though it commanded a minority of the votes of the country and a minority of the seats in Parliament, was called upon, as the most powerful single political group, to form the ministry and conduct the affairs of the nation. This position of precarious dominance was lost in the election in November of 1924, but the party increased its poll of votes and was recognized as the official Opposition of the Conservatives, who were then in power. After the election of 1928, the Labor Party, though still in the minority as compared with the combined strength of

the two opposing parties, again took command of the government. An election in 1931, precipitated by the national financial emergency, split the Labor Party into two factions, one of which joined with the Conservatives to form a non-partisan National Government, which has held office to the present time. This conflict has weakened the political power of organized labor. In the political branch of its activities the British labor movement is avowedly socialistic, relying however upon peaceful and constitutional means to effect its ultimate aim of thorough social reconstruction.

History of the American labor movement: Before 1850. The labor movement in the United States was both later in origin and of slower growth than in Great Britain. The industrial revolution, which supplied the conditions indispensable to the formation of labor organizations, was delayed by the predominantly agricultural character of economic life and by the institutions of the frontier, which prevented the formation of a stable wage-earning class. No unions of the modern type existed during the colonial period. Shortly after the adoption of the Constitution the industrial life of some of the larger cities began to show signs of social division between wage earner and employer and the beginnings of group action among the former. Small and ephemeral unions developed among the skilled craftsmen of New York, Philadelphia, and Baltimore, which, having for their primary object the provision of benefits for the sick, devoted a part of their effort to improving their working conditions. These sporadic movements were a feature of city life throughout the first quarter of the nineteenth century, but they resulted in no permanent labor organizations, and in no case did the union grow to sufficient size to become an influential factor in industry. Between 1825 and 1850 a more vigorous trade union movement arose, with its centre in those cities where the wage-earning class was rapidly increasing. By 1836 one hundred and sixty unions, with a combined membership of 300,000, were in existence in the seaboard cities. Trade associations were formed within the cities, uniting the individual craft groups for common action, and attempts were made to bring together the geographical units of each trade into a national grouping. The movement was drawn into politics, both within the

cities and on a national scale; it also was misled into supporting many of the utopian and communistic experiments which featured this period. Weakened by its failure in these political and socialistic ventures and by the disastrous panic of 1837, all of the national and city-wide associations and most of the small local craft unions passed out of existence before 1850.

Local craft unions. Thus before the middle of the nineteenth century no permanent organizations of labor had made their appearance in the United States. After 1850 however the quickened pace of industrial development produced a labor movement which, despite decline in strength during years of panic and industrial depression, has continued to grow down to the present day. Like their predecessors, these unions were at first small combinations of craftsmen within a given trade and locality. Called into existence as experimental adjustments to the life problems of the workers and not in response to abstract notions of social organization, their forms of structure were adapted to the prevailing economic conditions of the country. At this time the typical industry was a small scale unit whose influence did not reach much beyond the limits of its own town or city. Moreover the working class was still divided for the most part into non-competing skilled groups, each performing a distinctive function and having few interests in common with other and different occupational groups, even within the limits of the locality. Within these narrow ranges the workers were thrown into competition with each other in their search for employment and thus tended to develop feelings of common interest in opposition to the interests of the local employers. The small, local, narrowly exclusive organizations of skilled craftsmen, which appeared to be the only type of union capable of surviving at this time, owed their vitality to the fact that they represented a better adjustment to the economic conditions of the age than did any broader or more comprehensive grouping.

National craft unions. However the social forces working within the country were destined soon to stimulate an organization of labor on a larger scale. Steam transportation broke down the geographical barriers which had divided the labor market into small self-contained units and broadened the area within which

craftsmen were brought into competition with each other. In response to this development, the local unions began to combine along craft lines into national organizations, each limiting its membership to workers in a specific trade or calling. By 1870 upwards of thirty skilled trades had succeeded in integrating their local units into unified organizations covering the entire national area. More powerful forces tended to hasten this line of development. Beginning about 1870, business enterprise entered upon that process of nation-wide integration which is so prominent a feature of the modern industrial world. Gigantic industries, formed by the consolidation of formerly competing units or by the expansion of an individual enterprise, grew to positions of dominance in many lines of productive activity. The workers, scattered far and wide over the country, were brought face to face either with a single employer or with a number of employers so small that they could easily form working arrangements governing the terms at which they offered employment to the wage earners. Problems of wages, hours, and working conditions could no longer be viewed by each local group of workers as problems applying to themselves alone; the area of common interest broadened to conform to the increased scale of the typical business enterprise. Stimulated by these changes in the underlying economic organization of the country, the tendency toward national organization within each trade or craft proceeded at a rapid pace. At the present time the nation-wide union of workers in each separate trade is the dominant form of labor organization, although, as we shall see, a broader grouping of workers uniting all the trades within each industry has been gaining momentum in very recent years. In 1937, out of 169 unions then in existence, 133 were organizations embracing a single craft or a narrow combination of closely related crafts.

Comprehensive organization. But the forces sketched above were tending toward a still more comprehensive grouping of the workers. With the growth of national unity and the widening range of government activity in the sphere of industry, many problems made their appearance which touched the welfare of the wage earners in many different trades simultaneously. Reacting

to these forces, the labor movement in the United States from 1870 down to the present day has striven to set up a comprehensive, class-wide organization with sufficient coherence to weld together the entire mass of wage earners. At first, during the seventies and eighties, the experiment was tried of discarding all lines of craft division and uniting the working people into a single centralized organization. For a time this movement, as measured in terms of numerical strength, was successful. In 1886 the Knights of Labor, as the centralized union was called, numbered 600,000 members. But the grouping proved too comprehensive to produce coherent and harmonious action. The heterogeneous character of its membership — skilled and unskilled, men and women, farmers, wage earners, and professional men were enrolled together on a plane of equality — drove the Knights of Labor to support broad social and political policies and to waste its energies in sympathetic strikes and in ventures in the field of coöperative production which ended disastrously. Internal dissension arose within its ranks, and the organization began to disintegrate before 1890.

Its decline in power was accompanied by the rise of a somewhat different type of comprehensive union organization, the American Federation of Labor, which has dominated the labor movement in the United States since that time. As its name implies, this organization is a federation, not a centralized national structure. It recognizes the necessity of separate national unions in each of the occupational groups. These autonomous craft unions have command over their own policies and programs of action, but they unite with each other to form a loose federal organization, for the purpose of solving problems which they share in common and of promoting unity of action within the ranks of the organized labor class as a whole. This partial merging of the whole labor movement into a single organization, though it lacks the features of close integration that were characteristic of the Knights of Labor, is nevertheless an indication that the trend of social forces in modern society is toward the creation of a distinct class of wage earners with a consciousness of common interests. In 1935 a competing national organization, the Committee for Industrial Organization

(the C.I.O.), was formed in opposition to certain policies of the A. F. of L. and to foster the industrial as opposed to the craft form of unionism. Of the 169 national unions in existence in 1937 there were 103 which belonged to the A. F. of L., 31 were affiliated with the C.I.O., and 35 were independent of both organizations.

Types of union structure: The trade union. In point of structure, union organization is an adjustment to the prevailing economic conditions of the time. Since these underlying conditions do not long remain stable in our rapidly changing world and since they are never quite the same in all branches of industry within a country or in different countries at a given time, the structure of the labor movement is constantly changing with the passing years and is never uniform throughout the whole wage-earning population at any chosen date. It is possible however to distinguish two types of labor organization standing somewhat in contrast with each other, which have been illustrated in the preceding sketch of the historical development of the labor movement.

The first of these is the *craft* or *trade union*. This type attempts to bring together all the workers who perform the same specific function and who are therefore likely to compete with each other. For example bricklayers, carpenters, plasterers, lathers, painters, though they all belong to the same major branch of industry, the building trades, are organized into separate unions in each locality. Similar craft divisions are to be found among the workers on the railroads. The locomotive engineers, the firemen, the trainmen, and the conductors are united into self-contained groups distinct from each other. The craft union is firmly established as a type of organization in all branches of industry where skilled craftsmanship has survived the effects of machine technique. The local units of each craft have then united into national organizations in response to forces tending to broaden the range of the labor market.

In order to bring about coöperation between the different craft unions, reliance is placed in the device of *federation*. These federations are both regional and functional. As an example of regional federation we find in each industrial centre the local craft unions federated into a city council for the purpose of solving problems of

common interest in the locality. Similar federations within the area of each state attempt to bring about coherent action on legislative programs which fall within the power of the state legislature. Finally the American Federation of Labor draws the whole system together into a national grouping whose chief component units are the national craft unions. A functional federation takes the form of a loose affiliation of different crafts employed in the same general productive process. Thus the building trade workers, although allied primarily to their separate craft unions, are all drawn together into councils which attempt to promote inter-craft coöperation on common problems. Similar functional federations are found in the printing trades, the metal trades, and on the railroads.

The industrial union. The second major type of union structure is the *industrial union.* This form attempts to obliterate craft and trade divisions among the workers and to make membership in the union coterminous with the limits of each separate industry. From what the reader now knows of the economic organization of the modern world, he will understand how the factory system, with its machine technique and its refined division of labor, has tended to break down craftsmanship and undermine the position of the skilled worker. This process has been accompanied in the United States by a mounting tide of immigration, which has flooded the labor market with alien unskilled workers, whose traditional standard of living reduces the plane of competition to a level destructive of the higher standards of the native born artisans. When skilled and unskilled, native and alien, male and female struggle together to obtain employment, older lines of division between occupational groups tend to disappear, and the whole group of wage earners in each major line of industry tends to develop a consciousness of common interests. Organization along craft lines is out of place in this setting; adjustment to the conditions of the workers' life calls for a type of organization which will bring the workers in each branch of industry into a single union regardless of differences in the specific functions performed.

A prominent example is to be found in the mining industry. The United Mine Workers of America, in its local branches, brings

together all the workers who have any connection with the production of coal. Stationary engineers, carpenters, draymen, unskilled laborers, as well as the actual miners of coal, are united in a single organization, and the local units are merged in a national industrial union which aims to cover the entire mining area of the country. So it is in the garment making trades, where the Amalgamated Clothing Workers represent a single union of wage earners engaged in all the different processes involved in the production of clothing. The industrial union, like the craft union, consolidates its local units into national organizations, which attempt to perfect a monopoly of labor in each branch of industry throughout the country as a whole. In 1932 there were 32 of these industrial unions in the country. Some of these national industrial unions have joined the A. F. of L. in the desire to coöperate with the national craft unions in the solution of problems affecting the whole class of wage earners; but their presence in this organization is an exception to the prevailing craft unionism upon which the Federation is founded.

The C.I.O. The aggressive organization drive of the Committee for Industrial Organization, launched in 1935, caused a rapid growth of industrial unions, chiefly in enterprises formerly unorganized, and precipitated a bitter fight with the A. F. of L. In origin the C.I.O. was merely a working alliance among the leading industrial unions of the time, the United Mine Workers, the United Textile Workers, the unions in the clothing trades, and others. The Committee was formed primarily to attempt the organization of the large mass-production industries, such as steel, rubber and automobiles, which were almost wholly non-union. It fell afoul of the A. F. of L. almost at once, because of its outspoken condemnation of the Federation leaders for their apathy toward unionizing these industries and because the C.I.O. favored the industrial form of unionism, which brought it into conflict with certain craft unions, affiliated with the A. F. of L., which claimed jurisdiction over some of the occupations in these industries. Despite the active opposition of the A. F. of L., the C.I.O. pressed its organization drive vigorously in 1936 and 1937, with dramatic results. Its aggressive tactics precipitated occurrences in the

business world — the "sit-down strike," mass picketing, violent conflict with the police and the company guards — which aroused widespread national interest. We shall refer to these phases of the movement in the next chapter. It is in point here to note the success of the C.I.O. in its organizing campaign. Early in 1937 it won the recognition of the United States Steel Corporation, traditionally the leading anti-union concern of the country, in the form of a collective agreement covering all divisions of the corporation's enterprise. By the end of 1937, although it had lost bitter strikes with several large independent steel companies, agreements had been formed with some 450 companies in the metal industries, embracing a half-million workers. In the automobile business the new industrial unions won recognition from all the large companies except Ford. Similarly successful were the efforts to organize the rubber and electrical industries. The outstanding accomplishment of this movement was to establish unionism for the first time among unskilled and semi-skilled workers in the large-scale enterprises of the country.

As stated above, the C.I.O. was in origin only a joint standing committee of certain large unions, mostly within the A. F. of L. It still retains a technical affiliation with the A. F. of L., although ten of the unions have been expelled for insubordination to the Federation's officers. But it has begun to construct a national organization, parallel to and independent of the A. F. of L., looking toward eventual complete separation. This has its basis in local industrial union councils, which draw together the different units of a city, with a superstructure of regional and state councils formed to bring about coöperation over wider areas. At the present time industrial unionism is thus organized on a state-wide basis in four states, and many local councils exist in other parts of the country. If this growth continues, organized labor in the country will presently be formed into two competing national structures typifying a sharp distinction between craft and industrial unionism. Beginning with an original membership of about one million, the C.I.O. now claims three and a half million members.

Extent of unionism in the United States. The following table shows the growth of unionism in this country since 1900.

UNION MEMBERSHIP IN THE UNITED STATES

Year	A. F. of L. unions	All unions
1900	625,200	868,500
1910	1,630,300	2,184,200
1920	4,156,000	5,110,000
1930	2,961,096	4,100,000
1937	3,268,600	7,296,500

The figures for 1937 show the disruption of the union movement through the recent growth of industrial unionism under the C.I.O. Of the four million members at present outside the A. F. of L. the C.I.O. claims three and a half millions; the remainder are in unions not affiliated with either federation At the present time (1938) the union movement is growing rapidly under the favoring influence of "New Deal" policy. It is however too early to say whether these recent recruits to the unions will continue as a permanent membership.

Political activity of American unions: Politics generally avoided. In one regard the American labor movement has not followed the lines of development marked out by the history of the older movements of European countries. During the past three decades, in which period the unions in the United States have made their most substantial and continuous gains in strength, they have abstained from organizing an independent labor party. Most of the larger national craft unions, as well as the A. F. of L., have incorporated in their constitutions provisions which forbid their members from attempting to resolve the organization into a political party. With the recent growth of industrial unionism there has been some departure from this practice, the C.I.O. having entered the political arena in certain important industrial areas, sometimes with its own candidates for office. Nevertheless it may still be said that on the whole American unions conceive their problem to be primarily economic rather than political; their objective is to enhance their bargaining power as against the employers and not to capture the power of the state in the effort to put through far-reaching reforms. Accordingly their members are urged to retain their independence of choice of party membership, the official action of the unions in the way of influencing this choice consisting only

in their endorsement of those candidates presented by the major parties who have shown themselves "friendly to the interests of labor." The unions have exerted political pressure chiefly by this nonpartisan endorsement of candidates and by maintaining strong lobbies at the state and federal capitols.

This development is to be attributed to two major causes. Several of the early attempts to organize labor in America were destroyed by ill advised campaigns in politics, and the unions have learned to fear the divisive influence of partisan politics upon their membership. Furthermore the American governmental structure is such as to discourage independent activity on the part of any economic group. The division of sovereignty between national and state governments, with the important power over social legislation distributed among the forty-eight different commonwealths, the separation of the governmental machine into semi-independent executive, legislative, and judicial units, the power of a Supreme Court, remote from political influences, to pass upon the constitutionality of state and federal statutes, these and other characteristics of the system multiply the obstacles in the way of effective political action on the part of any minority economic group operating independently. Moreover the predominately *laissez faire* character of governmental policy has tended, at least until the advent of the "New Deal," to minimize the importance of legislation as a means toward the accomplishment of economic ends and has encouraged workers and employers alike to rely upon their own bargaining power.

Possibility of increased political activity. However the labor movement in the United States does not conform to uniform pattern with regard to this question of political activity any more than in other matters of structure and policy. A few of the strong unions have consistently asserted the necessity of carrying on independent political action along lines similar to those followed by the labor movement of England.

It is noteworthy that the unions which sponsor the formation of an independent labor party in this country are almost wholly unions of the industrial type, whose broad inclusive basis of membership compels them to adopt policies of general social import.

The attitude of the C.I.O. has been noted. The tendencies which further this form of unionism therefore will probably promote the political activity of organized labor.

Other tendencies may be observed working in the same direction. In all industrial countries governments are becoming an increasingly important force in economic affairs, regulating business activities, establishing standards of economic conduct, restricting individual initiative along many lines, entering the industrial world as owners of productive enterprises and as employers of labor. These expanding economic functions of government may attain such a degree of development that organized labor can no longer accomplish its aims by economic action unsupported by its organized political power, in which case we may expect the unions to organize for political action.

Communist unions. One group of American unions provides a striking exception to the general reluctance of the workers to enter the political arena as a separately organized party. These are the recently formed communistic organizations. The attempt of the Communist Party to draw the American labor movement into support of its revolutionary political program has gone through various stages of development. From 1920 to 1929 the Communists in this country, acting on instructions from the Red International at Moscow, attempted to spread their propaganda within the established unions rather than by forming separate organizations of their own. The vigorous expulsion of these revolutionaries by the existing unions, most of which are essentially conservative, led to a change of tactics in 1929. Again acting on instructions from Moscow, the Communists then formed their own unions paralleling those of the conservative type. By 1934 there were ten of these unions, all affiliated in a national organization known as the Trade Union Unity League. In that year their lack of success in dividing the labor movement caused the Communists to return to the former tactics of "boring from within." They took an active part in the C.I.O. drive and succeeded in forming radical minorities in many of the new industrial unions. At the present time (1938) the C.I.O. is in process of expelling these communist factions from its membership.

Legal position of unionism in America: General situation. The activities of trade unions have frequently been condemned by American courts,[1] but at no time since their origin have unions as such been illegal in this country. The law of the land permits voluntary organization of individuals for the purpose of promoting their interests, and this permission applies to wage earners as to other interest groups. Accordingly the right to organize and to act collectively has never been challenged by the legal system, provided the means used to effect organization and to carry out union policies have not been unlawful in themselves. Despite this legal immunity the right of organization has been fought by hostile employers with methods which have at times made organization virtually impossible.

The most formidable of these methods has been the device known in trade union circles as the "yellow dog" contract. This was an agreement entered into between the workers and their employers which bound the workers as a condition of employment to abstain from joining any trade union. The obvious effect of such an agreement was to punish with discharge any wage earner who subsequently decided to become a union member. But the "yellow dog" contract did more than this; it was construed by the courts in such a way as to forbid the union from making any attempt to organize the workers in industries covered by it, on the ground that such attempt was incitement to breach of contract. Until recent years this anti-union device was much used by employers' associations, particularly in the mining industry, thus creating areas of industry within which the right to organize was, as a practical matter, denied. The national industrial recovery act of 1933 specifically forbade the "yellow dog" contract. Furthermore the act acknowledged the unequivocal right of labor to bargain collectively through representatives of their own choosing. To give positive encouragement to the growth of unions, the business codes set up under the National Recovery Administration were required to contain provisions for collective bargaining in every enterprise embraced by them. It is true that the mandatory code system

[1] This topic will be elaborated in the next chapter, which deals with the industrial conflict.

was later declared unconstitutional, but the outlawing of the "yellow dog" contract and the general approval of collective bargaining remain as part of the national law.

National labor relations act. With the adoption of the national labor relations act in 1935, the federal government altered the attitude of American national law from merely permitting to positively approving and fostering union organization. This law lists a number of "unfair practices" of employers which are prohibited, including in the list many of the devices formerly used to prevent or retard union growth within the enterprise or to establish employer domination over the union when formed. Refusal to negotiate with the accredited representatives of an organization among the concern's employees is forbidden as an "unfair practice." The law creates a permanent administrative agency, the National Labor Relations Board, made up ostensibly of members not partisan to either side of the industrial dispute, together with subordinate boards covering the principal industrial regions of the country. Disregarding certain powers of the NLRB not pertinent to our immediate topic, it is to be noted that the board has taken aggressive action in favor of unionism. It employs the "unfair practices" sections of the law for this purpose, having been given broad authority to investigate, hold hearings, and make decisions of a judicial character whenever labor leaders in a given industry complain of "obstructive" tactics by the employer. In cases where the right of union officers to represent the employees is in dispute, or where there is disputed jurisdiction between independent unions and a company union, the board is empowered to take a vote of the wage earners, who thereby decide whether they wish to be organized and if so in what form of union. The outcome of such an election settles the question whether a given group of labor leaders is the accredited representative of the employees If this question is settled in favor of the union, refusal to recognize the union's officers as a bargaining agency becomes an "unfair practice."

The activities of the board were at first retarded by doubts as to the constitutionality of the act, a question which was not set at rest until April, 1937. Down to July, 1938, the board had held over twelve hundred elections in different concerns, most of them within

the year immediately preceding, for the purpose of determining the desires of the employees with regard to unionism. It was successful in obtaining union recognition in many of these cases. In enforcing the "unfair practices" section of the law, the board obtained the reëmployment of some 200,000 workers after they had been on strike; it also required the reinstatement of 10,000 individuals in different concerns held to have been discharged for union activity. The national labor relations act applies to industries (aside from transportation) whose business operations cross state lines, which means of course all the large-scale industries of the country. In 1937 the great industrial areas of Massachusetts, New York, Pennsylvania, and Wisconsin, passed similar laws applicable to intrastate industry.

Characteristics of the independent union. The unionism we have been discussing is an independent and exclusive movement of the wage earners, stimulated in its growth and activities by a consciousness of common interests among the workers and of common opposition to certain interests of the employers. It is independent in the sense that it does not rely for leadership upon any other social class but supplies its leaders from its own wage-earning members and shapes its policies by consultation within its own group. Its exclusive character is shown by its refusal to admit to membership anyone identified with the propertied or employing classes. If a member of a union is promoted to a managerial position by his employer he must withdraw from the union. The movement contemplates a horizontal division of society into two opposing classes, which cut across the boundaries of the separate business establishments. It views the interests of these two classes as inevitably antagonistic and expects the two to organize separately for bargaining purposes or for industrial warfare when negotiation fails to settle the point at issue between them.

The company union. Another type of organization has developed in this country during the recent past, known as the "company union." This organization unites both managers and workers within each separate industry, presupposing mutuality of interest within the enterprise, and attempting by promoting harmony and efficiency to increase the competitive power of the business in its

struggle with other business units. The form taken by the company union varies in detail among different enterprises, but all such organizations have as their essential feature a system of joint councils made up of representatives elected by the wage earners, on the one side, and individuals appointed to represent the management, on the other. The jurisdiction of these councils also differs from industry to industry, but it is customary to give them some degree of control over business policies most closely affecting the welfare of the workers, such as the wage rate, the system of wage payment, disciplinary matters involving the discharge of workers, and the like. The joint council is also used as a medium to acquaint the workers with the general problems of the business and to enlist their interest in technical and organizational improvements which will promote productive efficiency.

Until the establishment by the federal government of the National Labor Relations Board the company union was invariably a creation of the employer. Frequently it was inspired by the employer's hostility to the independent union and his fear that the union movement would gain a foothold in his enterprise unless some gesture were made indicating his willingness to deal collectively with his own employees. Company unions of this sort are really a weapon of attack upon the labor movement and are so recognized by labor leaders. It is this sort of dependent organization that the National Labor Relations Board is attempting to suppress. In many cases however the employer is moved by a sincere desire to improve conditions within his business and finds no organization in existence to represent his employees. But whatever the motive which calls the company union into being, it is so opposed in social philosophy and in structure to independent unionism that conflict between the two is certain to arise whenever they come into contact with each other.

Growth of the company union. Isolated examples of the company union may be found prior to 1900; and between that date and the war period several large enterprises — among them the Colorado Fuel and Iron Company, the American Rolling Mills, the Philadelphia Rapid Transit Company, and the Hart, Schaffner and Marx Company — adopted this method of dealing with their

employees. Labor problems precipitated by the World War caused many other large concerns, such as the Standard Oil, Westinghouse, and Goodyear companies, to set up systems of joint councils as an efficiency measure. But it was during the years immediately following the war, when trade unionism was particularly aggressive in its attitude toward employers and began to grow in membership with unusual rapidity, that this competing type of organization attained really vigorous growth. At this time employers' associations launched a national movement in support of the company union as a form of industrial organization better suited to American traditions than the class-conscious trade unionism of the older type. The following table records the growth of company unions from this date.

GROWTH OF COMPANY UNIONS IN THE UNITED STATES
1919–1932 [1]

Year	*Number of concerns with company unions*	*Number of wage earners enrolled in company unions*
1919	145	403,765
1922	385	690,000
1924	421	1,240,705
1926	432	1,369,078
1928	399	1,547,766
1932	313	1,263,194

Although the number of business concerns employing this method of dealing with labor fluctuated from year to year — some companies abandoning the practice and others adopting it — the number of wage earners embraced in company unions grew steadily down to the crisis year 1929. It is significant that, whereas the independent unions lost over a half-million members between 1922 and 1928, the company unions gained over three-quarters of a million. Furthermore the decline in company union membership during the early years of the depression was less, both in absolute numbers and proportionately, than that suffered by the independent unions.

The same features of the NRA which promoted the growth of independent unions after 1933 caused a revival of company union-

[1] National Industrial Conference Board, *Collective Bargaining Through Employee Representation*, New York, 1933.

ism. Obliged by clause 7a of the national industrial recovery act to deal collectively with labor in order to qualify for a code, many large employers hastened to organize their workers into company unions, and others revived and extended moribund organizations already existing.

By 1934, according to one estimate,[1] there were 800 such organizations in the country, embracing three million workers. The C.I.O. drive against the industries in which the company unions flourished made use of these organizations of employees, frequently winning them over to support its demands on the employers. In some cases these industries signed agreements with the C.I.O. unions and the company unions went out of existence. In other cases, the independent and company unions remained in the enterprise side by side and fought each other bitterly over the right to function as the sole agency of collective bargaining.

The board created by the national labor relations act of 1935 has assisted independent unionism in its struggle with the company unions. Among the "unfair practices" prohibited by the act is any attempt on the part of the employer to dominate labor organization in his enterprise, even if the organization is a company union. Since most of the existing company unions were created by the employers the board has ruled that many of them are incompetent under the law and has called elections in the enterprises to determine whether the workers wished to be represented by the independent unions. These activities of the National Labor Relations Board have hastened the decline of company unions since 1936.

Non-economic activities of unions: Social features. In the following discussion we shall refer to independent labor organizations unless the company union is specifically mentioned. These activities are many and varied. They are not all economic in character and hence do not all fall within the scope of our present study. No permanent association can be formed among men which does not tend to take on functions apart from its major purposes; for these associations will be made the mechanism through which their members give effect to the many social impulses which are common to mankind, such as the desire for com-

[1] Daugherty, *op cit.*, p. 643.

panionship, for organized sport and amusement, for self-expression along many lines. As older community ties are broken down by the progress of industry, the labor union becomes, for many of the wage earners, one of the few remaining bonds uniting them to their fellows in durable association, and it is made to serve purposes which among other groups of the population are served by clubs and fraternal organizations. These incidental functions of the unions are important and should be taken into account in any attempt to understand the full significance of this movement.

Benefit systems. One of the oldest of the economic functions performed by labor unions has been that of providing the wage earners with a form of mutual insurance against some of the hazards of their lives. This function is discharged by what is called the "benefit system" of the unions. Most of the labor organizations of England and the United States have collected funds out of which to pay small incomes to their members in case of accident, sickness, or unemployment and to provide payments to the members' families in case of death. At the outset these benefits were financed by a crude system of assessment which called upon the members for contributions as need arose. In all the older and more firmly established unions however the system has now been placed on an actuarial basis similar to that which underlies all scientific forms of insurance. The members make monthly contributions into the treasuries of their organizations, the resulting funds being segregated from funds held for other purposes and kept invested in safe securities for the purpose of financing the benefits to which the contributors are entitled. The unions vary in the amount of emphasis placed upon this phase of their service to their membership. Some do not have the benefit system at all, and others discharge it in a desultory and unreliable fashion. But among the more powerful national unions it holds a prominent place and is responsible in no small measure for the success which these unions have had in building up and retaining their membership.

As examples of the importance of the insurance principle in the labor movement, the unions among the workers on the steam railways may be cited. A survey of the United States Bureau of Labor Statistics in 1928 showed that 78 of the national unions had

benefit systems in operation. In 1930 the unions affiliated with the American Federation of Labor paid out a total of some thirty-seven million dollars in benefits to their members. The benefit system is of particular importance in hazardous occupations whose risks the regular insurance companies are reluctant to assume. For example the Brotherhood of Locomotive Engineers has carried approximately two hundred million dollars' worth of insurance in force at a given time, has paid as much as two and a half million dollars in a single year, and has made total payments to its members since the system was instituted amounting to over sixty million dollars. The firemen's union has paid a total of over twenty-nine million dollars in benefits and now carries about one hundred and fifty million dollars' worth of insurance continuously. The trainmen's union has paid out more than fifty-nine million dollars since the benefit system went into effect.

The examples given will suffice to show the importance of this phase of union activity and to give some indication of the magnitude of the financial resource which the unions can wield by mobilizing the small individual contributions of their members. But important as is the function of mutual insurance in the lives of the workers, other organizations exist which discharge it quite as effectively as do the labor unions. Except in the case of the more hazardous occupations, insurance companies established as business enterprises offer the wage earners indemnity against industrial accident and disease; and modern states have undertaken to insure them against the greatest of their hazards — unemployment. At the present time the United States is experimenting with statewide insurance against unemployment, and the experience of England has shown that such insurance can be supplied through other mechanism than that provided by the unions. The unique and distinctive economic functions of the labor organizations therefore do not inhere in their benefit systems. They consist in the attempts of the unions to raise wages, shorten hours, and increase the economic security of the workers through the use of collective bargaining.

Collective bargaining: Purpose and uses. This device of organized labor has been examined in connection with our study of

the law of wages, which has made it clear that the workers may within certain limits raise their real wages by a change in bargaining methods which substitutes the collective for the individual wage contract.

We need pause here only to note that the bargaining superiority of the employer is, in special circumstances, increased by conditions peculiar to the locality or the enterprise. Such is the case for example when the workers are of foreign birth, ignorant of the language and industrial conditions of the country, and habituated to a low standard of living, or when they are women or children. Similar effects follow when the employer, in addition to his ability to supply employment, has other holds on the worker which he may use as weapons of intimidation, such for example as his ownership of all the available homes in the village or his position as creditor for arrears of payment due at the company store. We should note furthermore that the forces which make for the normal bargaining weakness of the worker grow in importance with the growth of the factory system. The larger the scale of the enterprise, the less competent is any individual worker to defend his interests.

Collective bargaining however is not of advantage to the laborers alone; it results in certain indirect benefits to the employer, which enhance its value in the industrial system. The mechanism through which this type of bargaining is carried out may be turned to account in drawing workers and employers together into closer contact on a basis of mutual tolerance and good feeling. Many of the wastes in modern industry and much of the friction and bitterness engendered in the field of industrial relations flow from the absence of such contact. Organization of the workers for collective bargaining and the establishment of a permanent structure for negotiation with the employers may effect a real improvement in the situation. The workers are supplied with a mechanism for voicing their grievances and are assured of a respectful hearing. Discipline in the shop is more effective when it bears on observance of the terms of a contract in the making of which the men have had equal voice. The subordinate officials are restrained from harsh or overbearing tactics; fair and tolerant treatment of the

workers becomes good policy in view of their organized power. Above all labor is given a recognized share in the control of those features of industrial policy which most vitally affect their welfare, and this tends to influence their attitude toward their jobs and toward the enterprise. It goes far toward supplying a substitute for that personal interest in the product and the fortunes of the concern which existed in the handicraft stage of economic evolution but was destroyed, with far-reaching results, by the growth of the factory system. These benefits from organized collective bargaining are now recognized by practically all modern employers whose enterprises have grown to any size. In England and other industrial nations of Europe and to an increasing extent in the United States as well it is very rarely that the large employer insists on dealing with his workers individually. The spread of the company union noted above shows how widely the principle of collective dealing has won recognition in non-unionized industries.

Conflict over the method of collective bargaining. But though the soundness of the principle of collective bargaining is generally conceded, there is no such agreement regarding the method of applying this principle. One distinctively American phase of the problem of industrial relations is the quarrel over the question of "recognizing the union." Organized labor insists that the workmen in each of the different enterprises must be represented in their bargaining by the regular officials of the union. These men are not on the payroll of any employer with whom they negotiate wage agreements but are professional bargainers representing all the members of their union within a given geographical area. Recognition of the union binds the employer to adopt this method of bargaining. On this point the issue is joined between the independent unions and employers adhering to the company union principle. These employers, though they are willing to deal with their workers collectively, insist that the workers' representatives must themselves be employees of the enterprise.

Strong arguments are presented in defense of each of these opposing systems of bargaining. The employers who refuse to recognize the union insist that the "outside" negotiator is ignorant of the peculiar circumstances of the business and therefore unable

to discuss terms adjusted to that business's specific needs. Moreover these professional union officials are frequently so aggressive and domineering that peaceful compromise is impossible. In many cases they have proved to be corrupt, willing to foment trouble for the sake of their own gain in power or money.

On the other hand, leaders of organized labor stress the weaknesses of the company union as a device for protecting the interests of the workers: (1) the individuals chosen to represent the workers lack courage because their jobs depend on the good-will of the employer, (2) these representatives, being ordinary wage earners continuously employed at their jobs, can acquire no knowledge of general trade conditions and no expertness in negotiation, (3) final control always rests with the management, which may veto the decision of the council and abolish the whole system at will, and (4) this method of organization binds the wage earners to their individual enterprise, subjecting them to its hazards and depriving them of the collective support of other workers in case they are driven to strike in self-defense.

The merits of collective bargaining. When the question is considered on its merits, divorced from other aspects of union policy, the case in favor of the union method of collective bargaining is conclusive. It is apparent, on the face of the matter, that the interests of the workmen cannot safely be entrusted to a representative who is dependent on the employer for his livelihood and who may be victimized by petty persecution of the foremen or discharged outright for diligence in the performance of his duty to the union. This method might in time easily result in the elimination of all representatives who were not sufficiently pliable to be acceptable to the employer, in which case the negotiations could not be called true bargaining. There is some inconsistency in the attitude of those employers who refuse to confer with union officials while yet acknowledging the validity of the collective bargaining principle. On their side they habitually carry on negotiations with labor through the service of managers who are employed for this purpose, not because they are stockholders or members of the ownership faction of the enterprise, but because they are experts of professional status. When the workmen designate a union official

as their representative they are imitating this established business practice.

However it is virtually impossible to consider this question solely on its merits, for the controversy over the method of collective bargaining touches the very life of the union itself and is colored by all sorts of considerations relating to the social utility of unionism. If the workmen were allowed to organize for collective bargaining only within the limits of each individual enterprise and were compelled to employ ordinary wage earners as their representatives, the union movement in the form of class-wide labor organization would lose its reason for existence. Hence the officialdom of the unions base their case in defense of their method of bargaining, not alone on its undoubted superiority over the alternative method, but also on broader grounds of the social desirability of the labor union itself. On the other hand, opposition to the union method of collective bargaining is most frequently inspired by antagonism to certain practices of organized labor and has the ulterior motive of destroying or enervating the union movement. The union device of the closed shop, in particular, is viewed as the inevitable and highly objectionable outcome of union recognition.

The closed shop. Before discussing the controversy which rages around the question of the closed shop, it is necessary to define our terms with some care. Strictly defined, a closed shop is one in which none but union men will be employed to perform those tasks whose labor supply has been unionized. Its opposite, the open shop, is one in which there is no discrimination in offering employment as between union and non-union men. Both of these terms are often ambiguously used. For example some employers who declare their adherence to the open shop principle announce that they will not hire union men at all. This policy obviously does not lead to a real open shop. Moreover the closed shop has been confused with the *closed union* in much discussion of its merits and defects. A closed union is one which refuses admission to properly qualified workers. The effects of the closed shop are obviously quite different according as this privilege is enjoyed by a closed union or a union open to all comers; but it is a mistake

to assume that the closed shop and closed union are necessarily found conjoined.

The closed shop is merely an extension of the union program of collective bargaining, with its implied corollary of standardization. Collective bargaining is based on the assumption that the resulting agreement will be binding on the entire labor force. If union and non-union workers are allowed on the same job in the same shop, the established standards might be disregarded by the employer in his dealings with the non-union workers. If this resulted in lower wages for some men than for others, it is obvious that preference would be given to the non-union workers whenever the two groups were competing for employment, and the purposes of collective bargaining would be thwarted. Hence from the point of view of the corporate interests of the union, there are obvious advantages in the closed shop policy. Clearly it sets at rest the question of union recognition. It enhances the worker's inducement to join the union and reduces the difficulty of expanding the field covered by organized labor. It strengthens the disciplinary powers of the union officials; the recalcitrant member expelled from the union suffers unemployment. This in turn facilitates the collection of union dues and makes for faithful observance of union rules. It is an absolute guarantee against the victimization of spokesmen for the workers when defending the interests of their constituents. It strengthens the hand of the union when, in the last resort, it orders a strike. In fact the closed shop makes the strike so formidable that the threat of the strike often suffices to win the point for which the union has been contending.

Advocates of the closed shop policy urge in its favor certain advantages to the employer. First it makes for peace in the shop. Union and non-union men can rarely work side by side in harmony; the loyal member of the union looks upon the non-union worker as a renegade, a traitor to his class. Secondly it gives the union power to enforce faithful observance of the collective agreement and to guarantee a standard of performance. Thirdly it insures the good-will of the buying public in the case of all those concerns whose products are sold chiefly among working classes; from this point of view the closed shop and the union label go

hand in hand. Finally it protects the "fair" employer against cutthroat competition from enterprises which mistreat their workers. This last argument holds good obviously only on the assumption that the closed shop applies to all units of a sharply competitive industry.

In the fight against the closed shop many arguments of a sentimental character have been used; such as that it infringes the inalienable right of every American to remain outside of the union, that it is contrary to the spirit of our political and social institutions, and the like. Ignoring these, it can be said that the economic arguments against the closed shop cohere about a common point; namely, that it places dangerous power in the hands of union officials and increases the menace of all forms of anti-social practice in which the unions are tempted to indulge. Much is made of the fact that the closed shop weakens the efforts of management to stimulate the efficiency of the working force. The employer's great disciplinary weapon is the right of discharge. It is asserted that the closed shop nullifies the effect of this weapon. Where this is true however it does not follow as a necessary consequence of the closed shop itself, but results from the indifference of the union to the efficiency of its members. If we could imagine that all unions were intent on supporting the efforts of management to reduce cost and save waste, it would be apparent that a closed shop policy would increase the disciplinary powers of the employer and act as a positive force on the side of efficiency. So it is with the other charges in the indictment of this policy; for example that it encourages the make-work practices of laborers, solidifies their opposition to labor-saving machinery, results in many monopolistic abuses. When we examine these charges, we find that they are directed against the uses to which the unions put their power, rather than against the specific device of the closed shop. Thus the final judgment on this phase of the union program must turn on the question of the general social philosophy of organized labor and the general economic behavior of unionized workers.

Restrictive tendencies of union policy. Thus far we have considered the economic program of organized labor merely as a collection of devices for increasing the bargaining power of the wage

earners. Collective bargaining and the closed shop are devices useful to labor in that they enable the workers to obtain the full reward for their services as set by economic law. They are all quite harmonious with an open union policy and in fact with the absence of any restrictive or monopolistic devices in the program of organized labor. But unless supported by such devices, their effectiveness to improve the economic condition of labor is strictly limited. The real income of the different groups of workers is controlled fundamentally by the numbers in these groups in relation to the demand for their services. An overcrowded group may make full use of collective bargaining, standardization, and the closed shop and yet, if all its members are to find employment, must be content with a miserably small income. Recognizing this fact, the strongly entrenched unions are at times prone to adopt various monopolistic practices.

These restrictive features of the union program fall into two major groups, each calculated to enrich the members of the union at the expense of some other class in society. In the first place, some unions adopt the practice of artificially restricting their membership This is done, not by a direct refusal to admit qualified workmen, but in certain indirect ways which reach the same result, for example by setting up prohibitive initiation fees or requiring an unreasonable period of apprenticeship. It should be clear that no union dares adopt such a policy until it has obtained a sufficient degree of competitive power. The closed union cannot exist under open shop conditions, for the reason that the members of the union would then be compelled to compete for employment in the same industries with the excluded workmen, whose competition would be the more destructive of union standards because they were unorganized. Also the right of collective bargaining through union officials must have been won before there can be any incentive to close the union; for obvious reasons the employers' method of applying the principle of collective bargaining is effective against the closed union. Collective bargaining and the closed shop, though innocuous in themselves, take on a sinister aspect when they are used as the indispensable first steps toward the creation of a monopoly of labor. By applying the policy of restricting its mem-

bership, an organized group may make of itself a short factor in industry, raising the value of its marginal man's product and hence the wages of all its members. These gains are made at the expense, in the first instance, of the excluded workmen who are forced to render service of a type which is more plentiful and of lower value than what they are qualified to perform. In the long run all members of society share the loss through the decline in the social income which must result from employing a part of the labor supply at relatively unproductive tasks.

The second type of restrictive practices consists in a variety of devices for reducing the normal output of union members. These have been alluded to in a previous chapter We have seen that most wage earners are disposed, under certain conditions, to restrict their output and adopt "make-work" policies. In large part this is a reaction against the menace of unemployment, in part also it is a deliberate attempt to enhance the value of the product by increasing its scarcity. Restriction of output therefore is not a unique policy of organized, as distinct from unorganized, labor, though it is true that the trade union is in position to apply the policy with greater effect. We may note too that these monopolistic devices are the stock in trade of all owners and managers of industry whenever they are profitable and it is practicable to adopt them. It must not be inferred that all unions practice the two forms of restriction which we have discussed. Some unions renounce their use on grounds of broad social welfare; others have not attained to sufficient strength to make them effective.

LII

THE INDUSTRIAL CONFLICT

The strike. Conflict in the field of industrial relations is forced upon the attention of the public when the bargaining system breaks down and employers and workers attempt to settle the issue by open warfare. Each of these groups holds a weapon in reserve to be used in the last resort; the employers possess the *lockout*, the workers, the *strike*. These are correlative weapons. The lockout occurs when an employer or employers' association withdraws employment from a group of workers pending the settlement of some issue which is in dispute. The strike is a collective withdrawal of labor by a group of workers until the employers consent to certain demands which have been made upon them. These weapons are alike in that their use has the effect of stopping production and throwing a burden of loss upon the public. But since the lockout is less frequently used than the strike, it does not receive the same amount of public attention.

Spontaneous quitting of work has always been the last resort of the wage earners when all other methods fail to win for them acceptable terms of employment. The unions have carried this weapon over into their own program of action, regularized its use, and made it more effective. All strikes involve loss to the workers; hence the weapon is not often used thoughtlessly but in general only after deliberation and in defense of what the workers conceive to be a vital issue. There are exceptions to this rule, particularly in America, where racketeering has frequently contaminated the labor movement; corrupt union officials will call strikes in their own personal interests or to victimize an enemy among employers or rival labor leaders. The menace of the strike has increased in direct ratio with the spread of unionism. Strikes occur much more frequently today than when the union movement was in its infancy. They take place on a wider scale and are more protracted and more bitterly waged, and each strike is more of a national calamity, embrac-

ing a wider section of the industrial area and paralyzing the sources of supply for goods essential to public health and safety.

Statistics of strikes in the United States. The magnitude of the strike problem will be made clear by a glance at the following table.

TREND OF STRIKES IN THE UNITED STATES, 1900-1937 [1]

Years	*Average per year*		*Average days lost per year*
	STRIKES	STRIKERS	
1901–1905	2,790	405,000	11,000,000
1916–1921	3,500	1,800,000	56,000,000
1922 1927	1,160	685,000	19,000,000
1928–1932	760	290,000	5,500,000
1933–1937	2,230	1,240,000	26,000,000
1919	3,630	4,160,000	140,000,000
1921	2,890	1,100,000	56,000,000
1929	920	290,000	5,300,000
1932	840	320,000	5,100,000
1937 (8 mos)	3,420	1,650,000	40,000,000

Strikes tend to increase in number and magnitude during prosperous years and to decline during depression. Periods of numerous strikes are also correlated with organization campaigns of the unions. These relations are shown by the figures given for individual years at the bottom of the table. The year 1919 was one of flourishing business activity, marked by an organization drive upon certain large-scale non-union industries; 1932, near the trough of the depression, was a year of comparative industrial peace; 1937 shows the effects of the C.I.O. campaign during the recovery pèriod immediately preceding; 1929, a prosperous year, is an exception, explained by the relative contentment of the workers and the conservatism of the union officers identified with the A. F. of L.

The inter-relation of strikes with the attempts of unions to expand their membership is of particular importance in American labor history. When the attempt of the union to organize the workers in non-union industries is fought by the employers, a strike is the inevitable result. The great steel strike of 1919, the shopcrafts strike on the railways in the early twenties, the steel and

[1] Adapted from Daugherty, *op. cit.*, p. 287.

motor industry strikes of 1936 and 1937 are examples. Certain unsavory aspects of the industrial conflict in this country — destruction of property, violence to persons — are also identified with these periods. There is evidence to indicate that the correlation of the strike evil with union growth has been due to the fact that pressure to introduce unionism has been generally resisted in this country and that such correlation does not necessarily reflect a preference for violent methods on the part of established unions. This inference is supported by the labor history of other countries where unionism is an accepted institution throughout the industrial world.

Sympathetic and general strikes. Sometimes workers strike who have no grievances of their own and who ask for themselves nothing in the way of better terms of employment. These are sympathetic strikes — strikes called by one group of workers for the sole purpose of aiding another group in a different business concern to win a struggle with their employer. Sympathetic strikes are usually limited to two types of situations. In the first place, the employer involved in the primary strike may be carrying on his industry by the use of strike-breakers, and union men everywhere are anxious to prevent him from marketing his goods. In this case the sympathetic strike is called to coerce another employer, sometimes in a distant city, into withdrawing his patronage from the concern in question. A variant of this type of sympathetic strike occurs when the first employer is not involved in a strike at all but is refusing to employ any but non-union labor in his concern. In the second place, if the primary strike is a fight to preserve the union from destruction, organized labor of the locality in quite unrelated industries or the other unionized crafts in the same industry may strike in support of a principle considered vital to all unionized workers. An example is the nation-wide strike of the railroad brotherhoods in 1922, called in sympathy with the strike of the workers in the shop-crafts who were struggling to win union recognition.

In its milder form the general strike is merely the result of a multiplication of sympathetic strikes within a localized area until all unionized labor in that area has stopped work. There have been

such general strikes in the United States, in each case limited to a single city, and all of them effective for a short time in bringing the entire economic activity of the city to a standstill. In its extreme form the general strike involves the whole country in an abrupt stoppage of work in all industries, including the public services. European and South American countries have experienced such nation-wide disturbances at various times. But the general strike in this form is a political, not an economic, weapon, employed usually in support of a revolutionary program to which the organized labor of the nation in question is committed.

It is not difficult for labor leaders to construct a theoretical defense of the sympathetic strike or even of the localized general strike. The interdependence of industry has developed to such a point that a strike in one business unit may be very seriously affected by what is done in other business units in distant localities. The loss of a strike called in defense of some vital principle of unionism may seriously weaken the position of unionized workers elsewhere. These considerations are especially persuasive when employers are united in nation-wide associations in opposition to the spread of unionism. However while subscribing to these ideas in the abstract, the officers of the American national unions do not favor the use of the strike in these secondary forms, for the very practical reason that sympathetic strikes are rarely successful and general strikes never so. The public at large reacts to a general strike as to an insufferable attack upon its security and safety and turns out *en masse* to crush the revolt. Similarly the employer suffering from a sympathetic strike has little difficulty in convincing the public that the strikers, having no grievances against his labor policy, are acting in an indefensible manner. These strikes do occur in the United States, but they are almost invariably called by local union officials against the advice of their national officers.

Violence in strikes. Every strike creates a situation in which violence to persons and property is very likely to develop, and many strikes have degenerated into a state of disorder closely bordering on civil war. It is noteworthy however that violence is normally associated either with an organization campaign in non-union industries or with strikes called to prevent the extinction of a union

when an employer who has formerly dealt with it has refused to continue recognition. In such cases the union is fighting for its life and the situation calls to the front the aggressive, bellicose type of leader rather than the statesman. There is an uncompromising attitude on both sides which creates a spirit more akin to war than to business negotiation. The widespread sit-down strikes of 1936 and 1937, which were tantamount to forcible seizure of the company's property, were a novel development of the war tactics of newly formed unions in such a situation. Most of the earlier instances of bloody and destructive conflict in American labor history — such as the Homestead Strike in 1892 — developed out of similar situations.

When violence occurs in strikes called by established and recognized unions in the course of bargaining over the terms of employment, the cause is usually the presence on the scene of conflict of non-union strike-breakers or the police force or militia. As pointed out in another place the free lance worker who accepts employment in an industry during a strike is viewed by the strikers as a "job-stealer," against whom methods of coercion may fairly be used when milder persuasive devices fail. As for the presence of the police or militia, this usually results from an appeal that property rights and the right of the non-striker to work are entitled to protection. Since these appeals always come from the employers and property holders, and never from the strikers and wage earners, the appearance of the militia on the scene of action is taken by the strikers as evidence that force is to be used to aid the employer to win his case. In addition to these major causes of violence, sporadic cases of disorder can be traced to the activities of small revolutionary factions among the wage earners. With some extreme radicals a strike is but a means to increase the bitterness of class conflict, and violent methods are merely devices to further this end. As has been stated in another place, the American unions are almost unanimous in their opposition to these revolutionary practices.

Though the number of strikes increases with the spread of unionism, the amount of violence in the average strike tends to decrease. This is due principally to two factors. In the first place, as a union solidifies its position and expands its membership it has less to fear

from the use of non-union labor in the rôle of strike-breaker. In the second place, peace on the strike front is a matter of the self-interest of the union. Violent strikes are rarely won by the workers. In the final analysis public sympathy is a determining factor in the outcome of strikes, and the public has no sympathy with violent methods. Knowing this and deciding on the strike policy after a careful calculation of the chances of success and failure, well-established unions generally take steps to suppress any violent or destructive tendencies of their members. Since union discipline varies directly with the degree to which they have perfected their organization, the expansion of the unions tends, for this reason as well as for the one just mentioned, toward a decrease of violence in the industrial conflict.

Legal status of the strike: In Great Britain. Long experience has shown that attempts to prohibit strikes by law, if unaccompanied by constructive provisions for the settlement of industrial disputes, are ineffective. During the first quarter of the eighteenth century, when unionism was condemned in England as conspiracy, all forms of strikes were viewed as evidence of criminal restraint of trade and were rigorously suppressed. Even after the acts of 1824-1825 had legalized the union, strikes were still condemned under the doctrine of conspiracy, and workmen engaging in them were punished. However the courts were unable to prevent the workers from striking; the blind opposition of law had the effect merely of inciting the strikers to violence, thus aggravating the evil. Legal prohibition had to give way before a public opinion generated by those economic forces which made the strike an indispensable weapon for wage earners under the factory system. With the removal of the conspiracy doctrine during the seventies, all forms of the strike have been legal in England except the general strike and strikes in certain public services vital to safety and health.

In America: Legislation and the courts. In the United States there is no single law governing the activities of the unions, since this is a matter lying within the province of the states. As regards strikes, state laws are so frequently changed and the opinions of state courts are so at variance with each other that it is impossible to phrase a general rule regarding the legality of the strike in this country. The

only approach to a general principle is this: that all strikes are legal, provided their object is clearly and solely to improve the economic condition of the strikers; *e.g.*, by raising their wages, shortening their hours, or bettering their conditions of work. Unless the strikers can prove that their whole intent is to benefit themselves, not to work injury upon others, this rule of law does not justify their conduct. In other words, the courts must look to the motives of the union and not to its objective actions for grounds upon which to decide as to the legality of a strike. Under the American law all combinations are criminal conspiracies when their purpose is to destroy property or to interfere with the rights of others. Since all strikes involve preconcerted planning and organization, it is not difficult to find in most of them some evidence of malice. As a result of their suspicion regarding motive, the state and federal courts have frequently attempted to solve the problem of the industrial conflict by the simple expedient of forbidding an open break between employer and organized employees, and they are still making this attempt. As a general thing the local primary strike usually survives the test of motive; but strikes on a wide scale, and especially the sympathetic and the general strike, are likely to be prohibited. Certain strike weapons are sometimes prohibited even in the case of legal strikes — for example, picketing — as working injury on innocent people not party to the dispute.

The injunction. The means most frequently employed by the courts in the United States to prohibit strikes is the injunction process. The injunction is an order of a court restraining conduct which threatens irreparable injury to property. Violation of an injunction is "contempt of court," punishable by fine or imprisonment at the discretion of the judge and without jury trial. This device of law has had a long history in Anglo-Saxon jurisprudence, being used as a protective device under a great variety of circumstances, but in no country except the United States is it used in connection with labor disputes. In America its principal use is by no means concerned with industrial conflicts, and even where so employed the injunction is more frequently applied to restrain specific acts of the strikers which threaten injury to property than to prohibit the strike itself. The injunction was first applied to labor disputes in

1888, and between 1900 and 1933 the courts used it with increasing frequency to prohibit union activities which were held to be of the nature of criminal conspiracies. In 1919 the injunction was used on a national scale in an effort to prevent a strike of the coal miners, and in 1922 a nation-wide strike on the railroads was enjoined by a federal court. In neither case was it effective.

These actions by the courts have aroused the bitter resentment of organized labor, which interprets them as evidence that the judiciary is controlled by the employers. Responding to union pressure, the federal lawmakers have made two attempts to prohibit by law the use of the injunction in labor disputes. The Clayton Act of 1914 undertook to absolve the unions from liability under the anti-trust laws and the common law doctrine of restraint of trade and also to control the use of injunctions in labor disputes. Subsequent interpretations of the law by the courts nullified these defenses. The Norris-LaGuardia act of 1933 dealt again with both these problems, prohibiting the enjoining of many specified activities, including most of the standard practices of unions. This law has not yet been tested in the courts. But whether or not the use of the court injunction in labor disputes is now obsolete, its inability to solve the industrial conflict seems to have been demonstrated.

Indeed the experience of the United States, like that of England, has shown that no solution to the industrial conflict lies along the line of mere legal prohibition. Strikes have been broken by the use of the injunction, and in some instances the employer has succeeded, with the aid of the court, in winning his case against the opposition of the workers. But the causes of conflict have remained unaffected. The workers cannot be compelled to surrender this weapon of self-defense so long as the compensating power of the employer to use the lockout is left undiminished, or so long as no alternative machinery exists for the adjustment of grievances.

Protecting the public against strikes: The public interest. So interdependent are the various parts of our industrial structure that no disturbance can occur in one field without affecting other groups of producers and the consuming public. An abrupt cessation of the flow of goods from any basic industry, such as mining, transportation, or the metal trades, may paralyze the whole pro-

ductive system and, if prolonged, cause grave suffering and hardship. As trade unionism has spread, widening the area embraced in a single industrial conflict, this realization of the public interest in the relations of employer and worker has impelled modern countries to experiment with various devices for the prevention or speedy settlement of industrial disputes. These devices may be grouped in an ascending order of increasing public intervention and coercion, with purely voluntary and informal arrangements at one end of the scale and the application of governmental compulsion at the other.

Conciliation. The least formal of the measures used to effect a speedy settlement of a trade dispute is *conciliation*. Strictly defined, conciliation denotes a method of settlement in which the employers and the workers, or representatives of these two groups, voluntarily meet together for discussion and agreement without aid or compulsion from the outside

Conciliation implies a high degree of organization on both sides and a habit of collective negotiation. It is a natural outgrowth of the development of trade unionism, on the one hand, and employers' associations, on the other. The method is best used for interpreting the meaning of doubtful points in an existing trade agreement or of working out a peaceful compromise between conflicting demands of an employers' association and a trade union, to be embodied in a future agreement. Such differences as these usually relate to specific, practical matters of detail, on which compromise is both possible and beneficial to both parties. In the absence of machinery which begins almost automatically to function for the settlement of these differences, the alternative is an open break and a trial of force, which causes an immediate loss to all concerned and leaves an aftermath of bitterness and resentment. The effectiveness of voluntary conciliation is not fully appreciated by the public, because in the nature of the case this process if successful forestalls a disturbance which might have attracted public attention. As a matter of fact, conciliation has a long history of orderly and successful operation in many highly organized industries.

Questions of general principle, such as the right to organize, the prerogatives of union officials, etc., cannot be settled by concilia-

tion. Agreements to bring such matters within the range of conciliation almost invariably cause a breakdown of the whole machinery of friendly negotiation. Furthermore experience has proved that it is necessary for the conciliation board to reach a unanimous decision on all matters referred to it. The board is made up of equal representatives of workers and employers; its proceedings are informal; its decisions cannot be enforced against a recalcitrant party. Unless decisions are unanimous, the constituents of those representatives who have voiced dissent are very likely to repudiate the verdict. A strike or lockout ensues, which brings the whole process of conciliation into disrepute.

Conciliation promoted by government. In many countries attempts have been made by the government to set up a permanent system of conciliation embracing the entire industrial field. The Whitby Council plan established by law in Great Britain in 1917 and the nation-wide conciliation system set up in Germany in 1919 and 1921 are examples. These programs assume thoroughgoing organization of labor and provide for a hierarchy of joint councils, composed of equal representatives of the trade unions and employers' associations, which will settle disputes in the individual plants, in the major regional divisions, and finally over the whole national area.

The adoption of such joint council plans has not been possible in the United States, because of the imperfect organization of the wage earners, and because of the refusal of many large employers to deal with the independent unions. The industrial conference called by President Wilson in 1919 to work out a national system along this line broke up when the representatives of employers refused to accept the principle of collective bargaining through union officials. Attempts of the federal government to encourage conciliation as a part of its plan for compulsory arbitration on the railroads have encountered similar obstacles. The Railroad Labor Board created by the transportation act of 1920 was intended to act in the last resort, after attempts had been made to settle disputes by joint conference between employers and unions, but the whole process broke down when the Pennsylvania Railroad refused to deal with any organization other than its own company union.

The conciliation machinery of the Watson-Parker act of 1926, which required railroads to negotiate with representatives of "labor's own choosing," was also disrupted by conflict between the independent and company union principles. The national and regional labor boards set up in 1933 and 1934 as a phase of the NRA were designed as machinery for mediation and arbitration, but they were based on the expectation that a wide structure of conciliation councils would be called into existence by the legal requirement of collective bargaining. In this case too the substructure of conciliation was disrupted in certain industries by the refusal of employers to recognize independent unions.

Governmental attempts to promote conciliation in America will continue to cause more industrial disturbance than industrial peace as long as the company union and the independent labor organization are in conflict with each other. At the present time, as noted in the preceding chapter, the boards set up by the National Labor Relations Act of 1935 and by subsequent state laws are promoting the spread of independent unionism at the expense of the company unions. This trend may lay broader foundations for conciliation.

Mediation. Mediation may be viewed as the first step in the direction of acknowledging the right of the public to a voice in the affairs of employer and employee. It involves a friendly and informal effort by some individual or agency not party to the dispute to bring the disputants together for the purpose of working out a peaceful settlement. The function of the mediator is diplomatic rather than judicial. That informal, non-coercive intervention of this sort is so frequently successful may be attributed to human nature. "When once a difference has been accentuated and, still more, when it has developed into an open conflict, both sides are apt to be striving for 'mastery' as well as for the particular object of the dispute. They stand to lose dignity as well as money; and, consequently, their obstinacy will be much greater than the material point alone can justify. Not only is this, as a matter of fact, so, but it is frequently known to be so by the parties themselves. They will often have considered some matter worth the *chance* of a rupture but not worth the certainty of one. Hence when the rupture actually arrives, all that is needed is some device for

facilitating withdrawal without undue loss of dignity from a position assumed for purposes of bluff. Even if, in the earlier stages of a conflict, no way seems acceptable, a point is sure to be reached sooner or later when one party will be willing to yield if it can 'save its face' in doing so. Hence the opportunity for the 'good offices' of the mediator." [1]

Mediators may be self-appointed men of prominence in the social or political world who come forward at the critical moment to volunteer their services. In England the prime minister of the day or some eminent man of noble rank has succeeded at times in mediating disputes when all attempts at direct negotiation had failed. In America mayors of the cities, governors of states, and even the President have performed this function with success. The eminence of such a mediator "flatters the disputants with a sense of their own importance" and inclines them to make concessions as a personal favor to him which they would refuse to make to an opponent.

Official mediation. Many countries have devised regular machinery for mediation in the form of standing boards of public officials. In the United States the Erdman act of 1898 provided that the Chairman of the Interstate Commerce Commission and the Commissioner of Labor should offer their services as mediators in disputes on the railways. Power to act as arbitrators was also given to these officials. Between 1898 and 1912 forty-eight applications for mediation and arbitration were made to this board, the majority of the cases being settled by mediation. When the Newlands act superseded the Erdman act in 1913, provision was made for a Commission of Mediation and Arbitration to be appointed by the President for a term of seven years. Of the seventy-one cases in which the services of the commission were employed between 1913 and 1917, most were settled by mediation.

The failure of the commission to settle a threatened national strike in 1917 marked the end of the usefulness of this agency. The law of 1913 which created the Department of Labor set up a mediation service under the Secretary of Labor, which has grown in size and scope of activity until it now employs seventy officials. Be-

[1] A. C. Pigou, *Economics of Welfare*, 1920, p. 386.

tween 1913 and 1931 the service acted as mediator in some ten thousand disputes, with success in over 70 per cent of the cases. Agencies created during the World War — the President's Mediation Commission of 1917 and the War Labor Board (1918–1919) under the chairmanship of ex-President Taft — were strikingly successful in preserving peace during that trying period. Mediation by members of the national and regional labor boards and by the Secretary of Labor as well as by officials of the regular mediation service proved again the usefulness of this device during the flare-up of industrial unrest in 1933 and 1934. Most of the states have added mediation to the list of duties of their industrial commissions or departments of labor.

At present the most active mediatory agency in the United States is the National Labor Relations Board, whose powers were described in the chapter immediately preceding. The board is concerned primarily with conflicts arising out of the recent rapid spread of unionism into formerly non-union territory. It has permission, though not the obligation, to mediate when such conflicts either have caused a strike or are threatening one, and it has used this power freely. Between April, 1935, and July, 1937, it procured amicable settlement of upwards of 6,000 disputes, some 1,200 of which had reached the point of a strike. Failing mediation, the board has power to issue mandatory orders of the "cease and desist" type, requiring the abandonment of practices which are held to be an unjustifiable cause of strife. In only 250 cases has the board felt obliged to use these powers.

Voluntary arbitration: In general. Next in the order of increasing public intervention is the voluntary submission of industrial disputes to an impartial umpire or arbitrator. Voluntary arbitration usually presupposes an existing trade agreement between organized labor and the employer which calls for the final determination of disputed points by a third party. Consent of both parties must be obtained before a cause can be submitted to an arbitrator, and the award is binding on neither party, since it is not enforcible at law. However when a trade agreement contains an arbitration clause, it is extremely difficult for either party to refuse later to submit a cause in dispute without loss of public support, and con-

tinued practice of arbitration gradually builds up a moral sanction which has great force in controlling conduct. The acceptance of the award is voluntary, but the very facts that all previous steps have also been voluntary and that the arbitrator has been chosen by both sides because of his impartial character give the award a sanction which is often more binding than a judgment based on force. Experience has shown that few unions or employers repudiate arbitration awards, provided the arbitrator has not overstepped the bounds of his mandate.

Unofficial arbitration agreements and procedure. The arbitrator, like the mediator, may be either a private individual or board or a public officer or public board. The typical unofficial machinery in America consists of a board of three, two representing employers and workers respectively, the third an impartial chairman selected by these representatives or named in the trade agreement. In such cases the decision almost invariably rests with the chairman. Yet the other members are not without their uses: they serve as channels of information to the arbitrator, and their presence on the board adds to the influence of the decision when it is announced. Many large industries give full-time employment to an impartial umpire with the duty of interpreting disputed points in the trade agreement and adjudicating small causes of controversy as they arise. The development of trade agreements between national unions and large employers or employers' associations has created a new profession of arbitrators, men skilled in the technicalities of the trade, whose position is akin to that of the judge interpreting law. The Hart, Schaffner & Marx agreement, covering a large section of the Chicago clothing market, provided for the services of such an official. The Amalgamated Clothing Workers, the United Garment Workers, the Journeymen Tailors Union, the Ladies' Garment Workers, and the United Cap and Hat Makers have so generally incorporated arbitration provisions in their trade agreements that approximately 400,000 workers in the clothing trades have been at one time embraced in this machinery for preserving industrial peace. In the printing trades, the International Typographical Union and the publishers recognizing the closed shop have a ten-year national agreement providing for arbitration of all disputes.

In the building trades these arbitration devices are local in nature, covering each city market independently; an example is the Joint Arbitration Board of Chicago. All of these agreements stipulate that there shall be no strike or lockout pending arbitration proceedings. It will be noted that this use of arbitration by prior agreement is feasible only in highly organized industries where trade-unionism and collective bargaining are established principles. All unions which are no longer fighting for recognition favor arbitration. The system is thoroughly established throughout the industrial world wherever trade-unionism flourishes.

Official arbitration. Almost all governments have set aside public officials or boards to act as arbitrators on request of the parties to a dispute, either as an incidental feature of their duties or as their sole function. Officials or boards empowered to offer their services as mediators may also in most cases serve as arbitrators. Virtually every one of the American states has made some such provision for voluntary arbitration. The Erdman and Newlands acts provided that the designated mediation officials were to act as arbitrators of disputes on the railways upon consent of both parties if conciliation failed to compose the controversy. The transportation act of 1920, the Watson-Parker act of 1926, and the amendment of this act in 1934 all attempted to strengthen the arbitration procedure of the railways. It is unnecessary to discuss in detail this vast array of state and federal laws.

On the whole these *ex officio* arbitrators have had little success either in this country or elsewhere. Both sides are inclined to distrust the impartiality and the capacity of men who owe their positions to political fortune. An improvement on the system is the appointment of a panel from which arbitrators may be chosen by the two parties, rather than the designation of a specific official. This is done in some twenty American states, whose laws provide for the appointment by the governor of representatives on the Board of Arbitration from nominees of employers and workers.

Results of voluntary arbitration. As long as the arbitrators do not much concern themselves with principles of abstract justice and right, voluntary arbitration has proved successful in promoting peace in industry. In simple cases, such as the interpretation of

doubtful clauses in an existing trade agreement, the arbitrator has little difficulty in reaching an acceptable decision by applying fair-mindedness and common sense to the problem. But when the controversy involves some change in the *status quo,* such as an increase or decrease in wages, alteration in hours, etc., the most effective procedure is for the arbitrator to adjust his award to the relative strength of the two parties. A shrewd appraisal of the power of the two organizations and of their determination to fight in defense of their positions gives the arbitrator a basis on which to forecast the probable outcome of a strike or lockout. He then attempts to phrase a decision which will anticipate this result. Wages, hours, and other industrial conditions fixed by arbitration are not deductions from economic law or from premises of public interest, but compromises between two contending forces. The document which promulgates the award is often embellished with reference to general economic principles, human welfare, costs of living, etc.; but these are usually intended to make the decision palatable to the defeated party and to win public support. The justification of the arbitrator's activities is not that it ensures the operation of justice and right, but that it saves employer, worker, and the public from the losses involved in industrial warfare, while yet according to the potential victor the spoils of war. Experienced private arbitrators will freely admit that they are guided by these "practical" considerations. Public boards and officials are prone to seek a higher plane on which to base their decisions, with the result that arbitration machinery frequently breaks down under their administration. It is not to be expected that an organized group will continue to favor this method of settlement if its net result is to deprive it of the advantages of a strong bargaining position.

Introduction of compulsion: Transition devices. There are several transition stages between such voluntary arrangements as have been described and a fully developed system of coercive public intervention. One intermediate stage is represented by laws which make submission of disputes to arbitration compulsory, while leaving the acceptance of the award voluntary. An illustration is the transportation act of 1920, creating the Railroad Labor Board, which has been discussed in another place. Another stage is reached

when attempts are made to give the award binding force after arbitration has occurred through mutual consent of the two parties. In an informal way this is sometimes done by requiring the parties to an arbitration agreement to deposit sums of money, which are forfeited in event of refusal to accept the decision of the arbitrator. The laws of about a dozen American states provide for punishment, by fine or imprisonment, of a recalcitrant party who has previously agreed to arbitration by public officials. These laws have proved impossible to enforce, and their presence on the statute books has acted as a deterrent to voluntary arbitration. It is impossible to attempt here even a summary of these various expedients which lie midway between voluntary and compulsory arbitration. But something must be said regarding one form of partial public coercion which has attracted much attention in the field of industrial relations; namely, the system of compulsory investigation.

Compulsory investigation: Experience in Canada and Colorado. The first announcement of the right of the public to knowledge of the facts surrounding industrial disputes and the first provision of machinery for the collection and publication of this information were contained in certain state laws in the United States. It is the experience of Canada however which has brought the device of compulsory investigation into most prominence.

A severe coal strike in the year 1906 led to passage of the Canadian industrial disputes act in the following year. The act applies only to public utilities, transportation agencies, and mines. It provides that any change in wages and hours must be preceded by thirty days' notice. If the change is unacceptable to one of the parties, no strike or lockout may occur until after investigation of the facts by a board appointed by the Minister of Labor and public announcement of a decision by the board. Fines may be imposed for illegal strikes or lockouts during the investigation period. After the announcement of the award however both sides are at liberty to resort to a trial of strength. Experience has shown that, though the Canadian act has not stopped strikes entirely, it has been successful in ninety-two per cent of the cases to which it has been applied. This record of success has been overshadowed by the fact that certain of the illegal strikes have been of large scale,

causing much loss to the public. The system has brought out one significant fact: that the coercive features of the law are much less effective than its voluntary features. In fact the attempts to impose legal penalties on recalcitrant workers have proved so futile that this provision of the law has become inoperative. During the first ten-year period only eleven persons were prosecuted, although upwards of 80,000 had broken the law, and only $1,660 were collected in fines as compared with $150,000,000, the maximum legally due.

Although twenty-five of the American states have conferred powers of compulsory investigation upon their departments of labor, only Colorado has given this procedure extensive trial. In this state a system very similar to the Canadian was set up by the industrial commissions act of 1915. This act also was passed at the end of a disastrous strike in the mines. The scope of the law has been altered at various times, but since 1923 it has been limited to the public utilities, intrastate railways, and mines. It copies from the Canadian law the requirement of thirty days' notice of change in working conditions, the prohibition of strikes pending investigation, and the public announcement of a non-mandatory award by the Industrial Commission. The period since 1915 has been on the whole one of industrial peace in Colorado; but it is impossible to say to what extent this was due to the operation of compulsory investigation and to what extent to a system of conciliation set up by the mine operators at the same time.

Appraisal of compulsory investigation. The strength of a system of compulsory investigation lies in the facility with which it arouses public opinion in support of that party to the dispute who agrees to accept the findings of the commission. Fear of public condemnation promotes a willingness to compromise which results in the settlement of many potential causes of strike by conciliation procedure, often without outside aid. Its weakness lies in the fact that it is almost universally condemned by organized labor. This is due principally to the fact that the enforced postponement of hostilities, which is the essential feature of the system, works to the detriment of labor and in favor of the employer. The success of a strike depends in large measure on the complete tie-up of the industry. A

waiting period gives time for the assembling of strike-breakers, the shifting of orders to other concerns, and similar expedients, which reduce the employer's loss. The strikers gain nothing by being compelled to wait and are in danger of losing the morale and *esprit de corps* which are vital to success in a protracted struggle. For this reason the system is viewed with suspicion by organized labor as a device to strengthen the hand of the employer. Irritating delays in the procedure of official commissions and a distrust of the fairness of mind shown by the ordinary political appointee are other causes of the opposition of labor unions.

Compulsory arbitration: The affirmative argument. Interference by the government in the relations of worker and employer reaches its extreme form when strikes and lockouts are entirely forbidden and provision is made for compulsory arbitration with an award enforced by governmental power. Many are convinced that compulsory arbitration of all trade disputes must be the inevitable outcome of the progress of law. Historically the growing interest of the state in the preservation of public order has led to a widening of the scope of law to cover actions of individuals and social groups which were previously considered purely private in nature. Revenge, the blood feud, and retaliation have been suppressed, as the state has gathered into its hands the duty of administering justice and protecting the private rights of its citizens. In advanced nations the settlement of disputes by private contest, in all phases of human relationship except the relation of employer and worker, has been discarded in favor of the judicial process sanctioned by the force of government. That the extension of law to this important field has been delayed is attributed by the proponents of compulsory arbitration to the fact that public interest in industrial relations has been of late development, arising only after the establishment of the factory system and the growth of trade-unionism. Recognition of the right of the public to a voice in the settlement of trade disputes, it is argued, must lead in time to a disarming of capital and labor for private warfare. This argument from analogy is impressive; but, as we shall see, it carries with it implications of grave social importance which are not always fully understood by those who favor compulsory arbitration.

Experience in New Zealand. New Zealand was the first country to adopt a system of compulsory arbitration. The first law was passed in 1894, after a period of bitter strikes. Since that date, except for the years 1932–1936, there has been an unbroken history of compulsion in New Zealand. The original system had as its basis a series of regional conciliation commissions covering the entire island, each commission under the direction of a salaried expert. Attempt was made to settle all grievances by voluntary adjustment with the aid of these commissions. Failing this, an appeal from one of the parties took the case before a Court of Arbitration, which had final jurisdiction over all disputes. Until 1914 strikes and lockouts were not illegal unless application had been made for an award by the court, but on that date all strikes were prohibited in the case of unions registered under the act. Registration by the unions was voluntary, and those which did not register were not denied the right to strike. However the ease with which an entire labor group could be registered by a minority of its members gave the system of compulsion wide sway. A law of 1913 set up a scheme of compulsory investigation, similar to that of Canada, applicable to unregistered workers. This law prohibited strikes during the investigation period.

The Court of Arbitration consisted of a president, taken from among the justices of the Supreme Court, and two other members chosen from nominees of the registered unions and the registered employers' associations respectively. The court had judicial power to compel the attendance of witnesses and the submission of books and documents. Violation of awards was punishable by fine, and each award held good for three years unless superseded by a new judgment of the court, even though the union to which it applied had cancelled its registration in the meantime. Awards issued in cases affecting a single enterprise could be made applicable to the entire industry, thus equalizing competitive conditions.

For reasons stated below the employers in New Zealand finally united against the law. Successful in a political campaign in 1932, they amended the arbitration procedure to remove most of its compulsory features, providing that, except for unions composed chiefly of women, submission to arbitration should require consent

of both sides of the controversy. In other words, New Zealand returned at that time to a system of conciliation with voluntary arbitration as a last resort. But the change was of short duration, for the election of 1936 brought a labor government into power, which reinstated the compulsory features of the law and widened its scope.

Compulsory arbitration of Kansas. Outside of Australasia one attempt has been made to apply the principle of compulsion. The state of Kansas in 1920 passed a law prohibiting strikes and lockouts in specified industries and creating a Court of Industrial Relations to settle all causes in dispute. The industries affected were those engaged in the manufacture of clothing and food products, mining, transportation, and public utilities. The court was authorized to intervene in disputes arising within these industries, either on its own motion, on appeal of one of the parties, on petition by a small group of taxpayers, or at the order of the attorney general of the state. Its powers extended to the fixing of wages, hours, and working conditions, and in order to prevent suspension of operations it was empowered to take over any enterprise, guaranteeing a fair return to capital and labor. Its awards were mandatory, sanctioned by penalties of fine or imprisonment. Early in 1925 however the court was shorn of most of its power by a decision of the Supreme Court of the United States, which declared unconstitutional the compulsory arbitration features of the law as applied to working time and by implication overthrew the whole process of compulsory arbitration whenever its effect was "to compel owner and employees to continue the business on terms not of their own making." Just prior to this decision the court had been abolished by the legislature, its powers being transferred to a public utilities commission. But the decision of the Supreme Court has indicated the apparent unconstitutionality of all measures seeking to apply the principle of compulsion to privately owned enterprises in this country, except in cases of emergency in industries clearly affected with public interest, as long as the federal constitution contains a clause guaranteeing "liberty of contract."

Judgment of the Australasian experience. Any inferences regarding the practical operation of compulsory arbitration must be

drawn from the experience of Australasia, where this method of promoting industrial peace has had thorough trial. For a time the world was inclined to accept the evidence as proof of success. New Zealand was known until 1905 as a "land without strikes." But the subsequent history of the system has shown that strikes cannot be prevented by compulsion. Since 1906 New Zealand has experienced frequent and bitterly contested struggles in the field of industrial relations. Australia has had a similar experience. The attempt of the Commonwealth to stiffen the penalty features of the law by imprisoning strikers precipitated such widespread revolt that the severity of the penalties had to be modified.

It is becoming apparent that the early success of these laws was due chiefly to the contentment of the workers with a system which operated continuously in their favor. The Courts of Arbitration favored the principle of unionism, even to the extent of compelling employers to accept the closed shop. Early awards resulted in continuous advance of wages and shortening of hours, thus conceding to labor larger material gains than could have been obtained from the process of bargaining alone. There were no strikes, because the causes of strikes were absent. The underlying reason for this result favorable to labor was the rapid and sustained material progress of Australasia during this early period of its industrial development, which made possible a continued increase of the wage income without encroaching on the necessary rewards of the other factors of production. When the rate of progress was retarded and it became impossible to grant the demands of the workers, rebellion against the system arose immediately. An epidemic of strikes broke out, which were the more bitter because they were levied in part against a government which was thought to be unfair to labor. When the workers used their political power to influence the process of arbitration, the result was a series of decisions which laid unsupportable costs on the business world. Owners of industry were then driven in their turn to take political action against the system, with the result, as we have seen, that it was deprived of its mandatory features. The political victory of labor which restored compulsory arbitration in 1936 added new emphasis to the fact that the state exercising such powers is destined to become the center of a struggle

between capital and labor. In the final outcome the political power of the contending parties, rather than economic principle, has been the deciding factor under New Zealand's system of compulsory arbitration.

Inherent defect of compulsory arbitration. This experience shows that the most serious defect of compulsory arbitration in a social order like ours is not its record of incomplete success but its essential incompatibility with the institutions of a capitalistic society. When a public body undertakes to fix wages, hours, and other terms of employment, it is ultimately compelled to discard economic principles in favor of a determination based on moral or ethical considerations. Free play of economic forces apportions an income to labor, as to the other productive factors, which measures its marginal indispensability in society's industrial structure. Boards of arbitration are forced by popular pressure to fix wages with the intent of guaranteeing a decent or wholesome standard of living to the workers regardless of their competitive weakness. No single concession to the workers will keep them content for any length of time. Like all other social groups, they press for a continued improvement in their economic status. The practical difficulty of enforcing the award puts pressure on arbitration officials to placate the labor side of each controversy. Penalties can be applied with ease to the ownership faction, whereas fines levied upon scattered thousands of workers cannot be collected without great trouble and expense. Consequently the standard of living used as a basis for wage settlements must be raised to keep pace with the expanding demands of labor.

But the determination of one type of income according to arbitrary standards of justice and right necessarily involves a similarly authoritative determination of other types of income. When wages are raised out of compassion with the workers or from motives of political expediency, a limitation is set on the shares of the social income distributed as interest and profits. A board of arbitration, setting out with the simple object of finding a practical adjustment of isolated disputes in individual enterprises, is confronted in time with the task of determining what is a "fair" rate of return for the infinite variety of services performed by land, capital, and manage-

ment, as well as by wage labor. There is no final solution of this problem in a society where reliance is placed on private initiative to supply requisite amounts of wealth and of labor and to undertake the direction of the productive process.

Considered from the standpoint of its essential economic implications, compulsory arbitration is correctly viewed as the opening wedge of socialism.[1] It involves a gradual interference with property rights, freedom of contract, and private management, whose ultimate consequence is not always fully understood by those who favor it. The next step after fixing a rate of wages which is not acceptable to the workers is to compel free men to perform labor at that rate. If interest returns are lowered to a degree which retards investment, the government is confronted with the necessity of maintaining and improving capital equipment. The strike weapon cannot be taken from the workers unless the employers are deprived of the correlative weapon of discharge. This and other forms of interference with the rights of management, coupled with a possible restriction of profits, may proceed to the point of driving private enterprisers out of industry. Such a contingency will place on the government the obligation of ownership and management of enterprises whose collapse would seriously affect public welfare. It must be admitted that these implications of a system of compulsory arbitration are as yet only theoretical; they have not materialized in the experience of the countries employing that system. But compulsory arbitration has been brought to a decisive test only within recent years, when the slower pace of material progress has forced into clear relief the conflict between what is desirable from the humanitarian viewpoint, as a reward for labor, and what is economically possible without injury to other interests in production. There is now a clear tendency in these countries which employ compulsory arbitration for the distribution of income to become an issue between opposing political parties which struggle for mastery in the machinery of arbitration.

The future of the industrial conflict. Though it is unsafe to venture a prediction regarding the future of a rapidly changing industrial order, certain conclusions may be drawn from the world's

[1] F. W. Taussig, *Principles of Economics*, 1921, Vol. II, p. 349.

experience with the various methods of solving industrial disputes which have been discussed. It is clear that organized labor cannot be forced to surrender its right to strike as long as the powers of the enterpriser remain uncontrolled. Public opinion will not support a system of industry in which men without power to resist are employed for private profit and under the command of the profit-maker. In England and America members of the police and fire departments of the cities have been denied the right of striking; but this is practicable because they are public employees who accept these forms of employment in preference to opportunities in other lines. The suggestion that an anti-strike clause be introduced into the transportation act of 1920 aroused such menacing opposition from labor that the bill had to be modified. It is a safe prediction that strikes cannot be prohibited even in such essential industries as transportation and mining unless the government is prepared to take over ownership and operation of these industries. This would mean that competitive control of prices, wages, and profits had been discarded in favor of some non-economic form of control.

No type of public intervention has as yet proved as effective in promoting industrial peace as have voluntary arrangements set up within industry by mutual consent of employer and worker. Conciliation, mediation, and voluntary arbitration are accomplishing much in the way of smoothing out trade disputes. Machinery for public intervention in the last resort is most useful when it is carefully designed to promote these informal devices for negotiation and settlement. Attempts to apply compulsion instantly arouse the suspicion and antagonism of labor. If they are persisted in, despite this antagonism, their ultimate effect is to force organized labor into politics, with a program of far-reaching social change which may mean the end of the present system of industry. But voluntary arrangements, though effective to a degree, have done little more than mitigate the menace of the strike. They do not go far toward removing the natural clash of interest which is inherent in the present relation between employers and wage earners.

SUGGESTIONS FOR FURTHER READING ON PART VIII, PROBLEMS OF LABOR

I. Population and Immigration

MALTHUS, T. R. *An Essay on the Principle of Population.* Sixth edition. London, 1926
THOMPSON, W. S. *Population: A Study in Malthusianism.* New York, 1915
WRIGHT, H. *Population.* London, 1923
DUBLIN, L. I. *Population Problems in the United States and Canada.* New York, 1926
CARR-SAUNDERS, A. M. *World Population.* New York, 1936
WILLCOX, W. F. *International Migrations.* New York, 1931
FAIRCHILD, H. P. *Immigration.* New York, 1925
JENKS, J. W., and LAUCK, W. J. *The Immigration Problem.* New York, 1926
STEPHENSON, G. M. *A History of American Immigration.* Boston, 1926
GARIS, R. L. *Immigration Restriction.* New York, 1927

II. Labor Problems: General Studies

DAUGHERTY, C. R. *Labor Problems in American Industry.* Boston, 1933
CUMMINS, E. E. *The Labor Problem in the United States.* New York, 1932
FURNISS, E. S., and GUILD, L. R. *Labor Problems.* Boston, 1925
FELDMAN, H. *Problems in Labor Relations.* New York, 1937

III. Unemployment

DOUGLAS, P. H., and DIRECTOR, A. *The Problem of Unemployment.* New York, 1931
BEVERIDGE, W. *Unemployment: A Problem of Industry.* New York, 1930
PIGOU, A. C. *Industrial Fluctuations.* New York, 1927
MITCHELL, W. C. *Business Cycles and Unemployment.* New York, 1923
FELDMAN, H. *Regularization of Employment.* New York, 1925
KEYNES, M. *General Theory of Employment.* New York, 1936
ROBINSON, J. *Essays in the Theory of Employment.* New York, 1937
Social Security Board. *Social Security in the United States.* Washington, 1937

IV. Accident, Hours, Child and Woman Labor

LAUCK, W. J., and SYDENSTRICKER, E. *Conditions of Labor in American Industries.* New York, 1917
DOWNEY, E. H. *Workman's Compensation.* New York, 1924
GOLDMARK, J. *Fatigue and Efficiency.* New York, 1912
WOODBURY, R. M. *Workers' Health and Safety.* New York, 1927
ABBOTT, E. *Women in Industry.* New York, 1913
HUTCHINSON, E. J. *Women's Wages.* New York, 1919
Publications of the Children's Bureau. United States Government, Washington, D. C.
COMMONS, J. R., and ANDREWS, J. B. *Principles of Labor Legislation.* New York, 1927
WOODBURY, R. M. *Social Insurance, An Economic Analysis.* New York, 1917

V. The Labor Movement

Harris, H. *American Labor.* New Haven, 1939
Perlman, S. *Theory of the Labor Movement.* New York, 1928
Lauck, W. J. *Political and Industrial Democracy, 1776–1926.* New York, 1926
Webb, B. and S. *Industrial Democracy,* Part III. New York, 1920
Wolman, L. *Growth of American Trade Unions, 1880–1923.* New York, 1924
Brooks, R. R. R. *When Labor Organizes.* New Haven, 1937
Wolman, L. *Ebb and Flow of Trade Unionism.* New York, 1936

VI. The Industrial Conflict

Fitch, J. A. *The Causes of Industrial Unrest.* New York, 1924
Mason, A. T. *Organized Labor and the Law.* Durham, 1925
Gilchrist, R. N. *Conciliation and Arbitration.* Calcutta, 1922
International Labor Office. *Conciliation and Arbitration.* Geneva, 1927
Selekmen, B. M. *Postponing Strikes.* New York, 1927
Witte, E. E. *The Government in Labor Disputes.* New York, 1932
Feis, H. *The Settlement of Wage Disputes.* New York, 1921
Yellen, S. *American Labor Struggles.* New York, 1936
Milne-Bailey, W. *Trade Unions and the State.* London, 1934

PART IX

PROPOSED CHANGES OF THE ECONOMIC SYSTEM

LIII

PROFIT SHARING AND COÖPERATION

There are many who question the soundness of the wage system as a form of industrial organization. Its failure to stimulate the efficiency of labor, the friction between employers and workers, and the failure of all devices to solve the problem of the industrial conflict have called into existence various programs which aim at a more or less complete reorganization of the economic order. The most extreme of these reform movements indict the underlying institutions of private property and profit-seeking enterprise and propose thoroughgoing reconstruction on socialistic principles. These revolutionary programs will be studied in the following chapter. Other proposals for modification are less radical, endeavoring merely to better conditions within the present institutional framework of society.

Profit sharing: History. The profit-sharing movement has flourished at various times during the past century, only to lapse into decline when conditions ceased to favor its growth. The striking success of this device in the Léclaire establishment in Paris, where it was introduced in 1838, and in the Godin Familistère, which adopted it two years later, attracted much attention among progressive employers in Europe and Great Britain. During the middle decades of the nineteenth century many experiments along this line were begun, and for a time profit sharing was hailed by those with a bent toward reform as a complete and final solution of the labor problem. This early enthusiasm was short-lived. With a few notable exceptions, the first profit-sharing schemes collapsed, leaving those who had pinned their faith to them disillusioned and others wary of the risk of failure.

A renewed interest in profit sharing developed in the United States in the present century, particularly after the World War, when the aggressiveness of the trade union movement gave employers an incentive to seek a means of increasing the loyalty of their

employees. At this time many large concerns adopted profit sharing as the basis of their labor relations, and the spokesmen for the movement again proclaimed its virtues as a solution of the industrial conflict. Before 1930 the movement had spent its force, and many of the recently adopted plans were abandoned; while during the depression years which followed the experiment declined rapidly. At the present time there are only a few score plans in operation in this country, and although some of these — notably the Procter and Gamble, the Endicott-Johnson, the Dennison, the Filene, and the N. O. Nelson plans — have had a long record of success, the movement as a whole has now little more than historical interest. Its history would indicate however that profit sharing may again revive under favorable conditions; and for this reason its merits and limitations should be understood by students of modern industrial problems.

Prerequisites of true profit sharing. The essential principle of profit sharing is simple. It introduces a new basis of distributing the corporate income. After normal returns to the ownership and labor factions have been paid, the residual profits (ordinarily accruing entirely to the owners) are divided between the two according to some predetermined scale of sharing. Many arrangements which are not true profit sharing parade under that name. Some employers for example have the habit of bestowing gratuities on their employees at Christmas time or bonuses at the end of a prosperous year. The purpose of these pseudo profit-sharing schemes may be the same as that of true profit sharing, though usually it is not. "Such methods are of questionable value viewed as instruments either for the promotion of efficiency or of social welfare. The arbitrary decision of the management, made after and not before the extra exertion, robs the plans of a large part of their stimulating power. It is not to be expected that men will concern themselves to increase profits and keep down costs, when they have no knowledge as to what portion of the profits, if any, they will receive. From the standpoint of social betterment, such arbitrary distributions are also to be condemned. A widespread use of profit sharing of this character would inevitably work undesirable results among members of the employed class, by making them more dependent

on the bounty and charity of employers than on their own initiative and bargaining power." [1]

True profit sharing must be free from any taint of charity or philanthropy, a business arrangement between voluntarily contracting parties. The plan must be definite and clearly understood in advance and must bestow on labor a right to automatic increases in income under certain conditions. These contingent earnings of labor must be in addition to the market rate of wages. Profit sharing will not be effective unless the workers embraced in the scheme are guaranteed an income equivalent to that of other workers of their class and are given also the chance of sharing in the prosperity of their enterprise. Moreover in order to avoid suspicion that it is being used as an anti-union device, the scheme must leave the workers free to join labor organizations without prejudice to their standing as claimants of profits.

Methods of sharing. The precise basis of sharing residual profit between workers and owners varies with different profit-sharing plans. Some such arrangement as the following is typical. After payment of costs including wages, five per cent interest is paid on the capital stock, and the remainder of the net profit is then divided between stockholders and workers in the ratio that total face value of capital bears to total wages paid during the year. In some concerns the residual income is divided in the ratio of total interest (rather than capital value) to total wages; this obviously increases the amount of profits accruing to labor. It is customary to subdivide labor's share of profits among the individual workmen in proportion to their normal wages. This is a recognition of the fact that market rates of wages are an index to the relative productive importance of different groups of workers.

With very few exceptions profit-sharing plans provide that the workers' share of the residual income must be reinvested in the enterprise, being distributed in stock or other certificates of ownership. This is a natural arrangement in view of the purposes of the profit-sharing system. A gradually accumulating share of ownership in the hands of the workers will tend more completely to merge their interest with that of the employer than would a mere increase

[1] A. W. Burritt, *et al. Profit Sharing, Its Principles and Practice*, 1918, p. 11.

in their annual money income. It gives them in time a double stake in the prosperity of the enterprise; they share in the guaranteed interest income as well as in the current residual profits.

This feature of the system has the ultimate effect of giving profit sharing a social significance quite different from that with which it sets out. Successful experiments in profit sharing have paid the workers a minimum share in the profits equal to ten per cent of their wages, and sometimes this share has risen as high as twenty-five or thirty per cent. These percentages of the total payroll of a large enterprise represent a considerable amount of money which, being reinvested year by year, may in time give labor a surprisingly large fraction of ownership. This tends to grow as the years pass, unless measures are taken to maintain the proportion of ownership in the hands of other types of investors. The absolute amount of profits accruing to the workers generally exceeds that accruing to the owners, because the annual payroll of the typical business enterprise is larger than its capital investment. If now both groups of sharers are obliged to invest their part of the profits in the company's stock, labor's investment will grow faster than that of other investors. This process may result at length in transferring control over the concern to the workers, thus causing the profit-sharing scheme to culminate in a system of producers' coöperation, a system which raises social problems of quite a different order. This result is not inevitable, since it requires that the original group of workers and their descendants remain permanently attached to the enterprise. But since one of the effects of profit sharing is to reduce the turnover of labor, this event is by no means improbable. We shall examine the relationship of profit sharing to co-partnership and to producers' coöperation in later sections of this chapter.

Judgment of profit sharing: No radical change. Any appraisal of the merits of profit sharing must grant at the outset that no far-reaching social changes are to be expected from this device. It will be adopted only on the initiative of the employer when it holds forth promise of material advantage to himself. The employer retains control over the plan and can discard it at will. Only when a profit-sharing scheme has ripened into a system of producers' coöperation does it cause any fundamental change in the wage

system. This eventuality, though always a potential feature of profit sharing, will never materialize unless the scheme is under the direction of an employer convinced of the undesirability of the wage system. Profit sharing must be viewed as a plan with definite and narrowly practical objectives. Its purpose in the modern world is twofold: to stimulate the productive efficiency of the workers, and to promote harmony of industrial relations. The latter is subordinate to the former of these purposes; the employer is interested in industrial harmony as a means toward greater efficiency.

Limited usefulness in promoting efficiency. As a device of management to promote efficiency, the scope of profit sharing is limited. In the nature of the case it can be effective only when the character of the occupation is such that a clear relationship exists between the workers' faithfulness and industry and the profits of the concern. But this relationship is by no means clear in the case of ordinary types of labor in a large business unit. The profit of the concern is strongly affected by factors quite beyond the workers' control. Even aside from these extraneous forces, the effect upon profits of greater industry on the part of the individual laborer may be impossible to detect. Refined division of labor makes the connection between the productive efficiency of a single wage earner and the profits of the enterprise remote and indirect.

The limitations of profit sharing as a device for promoting efficiency have been summed up in two formulas by a careful student of the modern experience.[1] First *the efficiency of profit sharing varies directly with the rank of the worker.* Second *the efficiency of profit sharing varies inversely with the size of the group.* The reason for the first of these conditions is not far to seek: the effort of each individual in the higher groups counts for much in the corporate income. As for the second condition, profit sharing relies on the public opinion of the group to promote efficiency. The slacker is an enemy, not only to his own financial interest, but to that of all other members of the profit-sharing scheme. The larger the group, the less alert it will be to detect delinquencies of the individual, and the less effective will be its reprobation. For this reason profit sharing among the rank and file of the wage earners has rarely been

[1] A. W. Burritt, *et. al.* *Profit Sharing, Its Principles and Practice,* 1918.

effective except in those industries where the labor force is broken up into small, closely coöperating gangs. Obviously these limitations greatly restrict the field within which profit sharing can operate as an effective method of labor management. It is equally clear that any scheme of economic reorganization which merely increases the incomes of prosperous middle-class workers and which is impracticable for business reasons in the case of the great mass of wage earners has little to offer from the standpoint of general social improvement.

Limitations as an agency of industrial peace. The second major purpose of modern profit-sharing plans is to solve the problem of industrial conflict by removing the clash of interest between employer and worker. The logic of the system on this point is clear; it creates a situation in which neither party to the wage contract can prosper without sharing his prosperity with the other. There is no doubt that profit sharing accomplishes this purpose as long as the enterprise flourishes and there are profits to divide. There are few workers whose contentment is not increased by bread and butter benefits. But the great weakness, from this point of view, is that the system seldom survives a period of business adversity and that its breakdown is likely to leave the relations of employer and worker more strained than if no such experiment had ever been tried. This is the more likely to the extent that the scheme has been successful in arousing the industry of the workers; for then they will feel that they have earned an increase of income and cannot justly be denied it. If the reverse comes at the end of a long period of success and if it is not too frequently repeated, the profit-sharing scheme may weather the storm. Otherwise it is likely to break down, leaving an aftermath of bitterness.

One special reason why profit sharing has but limited power to prevent industrial conflict is that trade unions are almost unanimous in their condemnation of the system. Their principal official argument against it is that it is merely a device for lowering the normal wage rate. The employer, they say, will never pay labor more than it earns; profit-sharing schemes permit the employer to withhold a portion of these just earnings and to assume a pose of virtue when delivering this portion to the worker under the guise of

"profits." The result will be a weakening of labor's insistence on a wage adequate to its economic worth. Behind these official pronouncements lies organized labor's real reason for opposing profit sharing, the fear that a spread of this system will disintegrate the labor movement. Employers are largely responsible for this attitude. Profit sharing has often been used, frankly or covertly, for the purpose of smashing the union within an industry. It is safe to say that it will never succeed in destroying the labor movement and that it cannot hope to improve industrial relations as long as it is used for this purpose.

Co-partnership: Description. This name is given to systems of industrial relations which provide favorable means for the acquisition by the workers of part ownership in the industry. We have seen that most profit-sharing schemes attain this result by providing that the share of profits going to the wage earners be reinvested in whole or in part, but co-partnership is not merely an aspect of profit sharing. The policy of encouraging their employees to become part owners of the enterprise has been adopted by many large American businesses which do not practice profit sharing. Needless to say, this policy is not merely a device to increase the financial power of the business. Most stock acquisition or co-partnership plans involve special inducements to investment by the wage earners, which make this method of obtaining new capital more expensive than through ordinary banking channels. The policy is intended to make returns indirectly through its effect upon the attitude of the employee toward the management and toward the business as a corporate enterprise. If the policy became sufficiently general to reach the great mass of wage earners, it would wipe out the clear lines of demarcation between labor and capital and thus remove a potent cause of class conflict in modern society.

Co-partnership plans vary so much in detail from company to company that no single scheme can be taken as typical. It is usual however for the company to offer its stock to its own employees at a lower price than rules in the market for the general purchaser and to provide for instalment buying on easy terms. Usually too some sort of guaranty is given the purchaser that he can resell the stock to the company without loss, or can borrow against it at low rates

of interest if need arises. Participation is usually open to workers of all grades who have been in the employ of the company for a certain length of time; though in some cases the lowest classes of wage earners are excluded, while in plans not carrying this restriction the customary provision that the eligible individual may subscribe for stock in proportion to his annual earnings acts in the same way to throw the major holdings of stock into the upper ranks of the working force. Special devices are frequently employed to encourage the participators to retain the stock even when the market price makes selling profitable. Long-continued service in the company is rewarded in various ways: by gifts of additional stock, by progressively improved terms of purchase, and similar devices.

Co-partnership in American industry. At least one plan of co-partnership — that of the Illinois Central Railroad — was established prior to 1900; and the experiment of the United States Steel Corporation has been in operation since 1903. But the really rapid spread of the scheme occurred in the decade following 1919. A study by the United States Bureau of Labor Statistics in 1916 showed less than fifty business concerns practicing co-partnership at that time; in 1926 another investigation recorded 233 plans in existence; while in 1928 the National Industrial Conference Board estimated that about a million workers in some three hundred companies were part owners in their respective enterprises, their combined holdings totaling over a billion dollars' worth of stock. By 1930 the list of companies with plans in operation included many of the largest and most firmly established enterprises; the United States Steel Corporation, Standard Oil of New Jersey, the American Telephone and Telegraph Company, the Pennsylvania Railroad, the International Harvester Company, the General Motors Corporation, Firestone Rubber Company are but a few examples. During the first year of depression these co-partnership plans suffered very little, according to findings of the National Industrial Conference Board in 1930. But two years later the situation had changed. The increasing distress of the wage earners had caused many of them to sell out their holdings of stock; several large concerns had abandoned their plans because of inability to live up to their agreement to repurchase the stock at guaranteed prices; other concerns had

changed the terms of the scheme to the disadvantage of their employee stock owners; no new experiments of this sort were being tried. It is noteworthy however that when the depression was at its worst the National Industrial Conference Board in another report reiterated its own confidence in the essential soundness of the system and represented the employers as determined to reëstablish co-partnership on an even larger scale when business conditions began to improve.

During the twenties, when co-partnership systems were spreading rapidly, there were many who hailed the development as the beginning of a new era in American industrial relations. The figures given of the number of workers involved and the number and value of shares held by workers are certainly impressive; and so too is the list of giant corporations which were committed to this practice on the eve of the depression. But careful analysis of the figures shows that the system, even at its maximum growth, was of minor importance in the labor world. In 1926 the Federal Trade Commission found that employees owned only 1.5 per cent of the total common stock in the country, and 2 per cent of the total preferred stock. Over half of the stock thus owned was in the hands of a minor fraction of the participators — members of the managerial, sales, and office staffs — who were really outside the social class of wage earners. Moreover, despite the large aggregate number of participators, only about 5 per cent of the wage-earning population were involved in these co-partnership schemes. The whole movement might be dismissed as of trivial social importance except for certain considerations which are advanced to prove that its potentialities have not received full demonstration. The period of extensive experiment was very brief; during this period the system did display amazing powers of growth; its growth was stopped by abnormal economic forces; it will be resumed when the depression is over. Holding in abeyance any judgment based on actual experience, we must appraise in general terms the merits of this plan of economic reorganization.

Weaknesses of co-partnership. Stock acquisition plans contain certain practical dangers both to the business and to the wage earner. From the business standpoint, since the dominant motive

behind most of the stock selling plans is not financial, they may lead to the investment of new capital in the enterprise in excess of the opportunity for its productive use. But more important is the possible adverse effect of such plans on the economic welfare of the worker. The average wage earner is in no position to assume risk of loss of his slender savings. Investment in industrial stocks, subject to inevitable risks of fluctuation in value, is unwise for any man whose small savings constitute his only safeguard against a "rainy day." The bonds of the company might offer a more appropriate form of investment for the wage earner; but corporations cannot increase their bonds without multiplying their fixed charges, thus enhancing the danger of turning the co-partnership plan into a disastrous financial venture. Furthermore the essential purpose of the plan is to share control with the workers, and for this purpose stocks are of course a more appropriate medium than bonds.

The financial insecurity of the plan from the wage earner's standpoint becomes apparent in times of depression. The collapse of the stock market coincides with increasing unemployment. Workers involved in stock purchase schemes are then in danger of losing their wage income and their savings at the same time; and the employer may be too absorbed in the struggle to save himself from insolvency to come to their rescue. During the recent depression following the crisis of 1929 many thousands of wage earners experienced the disastrous effects of these hazards.

Aside from these considerations of a practical nature, the co-partnership program may be criticized in terms of underlying purpose. It binds the wage earner too closely to the specific enterprise, thus reducing his mobility and his bargaining power. It is designed in part to reduce the turnover of labor; this is one purpose of the deferred payment feature and the provision of special benefits to those who remain continuously in the company's employ. Within limits this is advantageous to all, for labor turnover is a wasteful practice. But it may be carried to a point injurious to the wage earner. Having an investment at stake, he is less likely to make the most of his opportunities to improve his economic position by finding a job elsewhere or by bargaining to raise his wages in his present job. There is danger that he will remain attached to an

industry whose fortunes are on the decline. Moreover the plan may be used intentionally by the employer to weaken the resistance of the wage earner to unjust conditions of employment. If it results in protecting the employer from competition of other industries for the services of his employees, just treatment of the wage earner will depend too much upon the employer's high-mindedness and sense of fair play. Bargaining between individuals, one of whom has some hold upon the other, rarely works out well for the weaker party.

This danger will be lessened if the workers continue to be members of the trade unions. But these schemes are usually destructive of this relationship within the wage-earning class, often intentionally so. They are sometimes so constructed as to penalize the strike and thus further to weaken the resistance of the workers. Such is obviously the purpose of the provision of many plans which promises special favors to those who remain "continuously" at work and are "considered worthy" by the management.

The fact is that, in our present system of industry, no reliable substitute has been found for the worker's own bargaining power as a means of protecting his interests. Profit-sharing and copartnership plans may do something to promote industrial harmony and efficiency if used sincerely for these purposes, and if they do not interfere with the wage earner's freedom to combine with others to increase his bargaining power or to strike as a last resort in self-defense. To the extent that they encroach on these powers of labor they will fail of their other purposes by ultimately arousing insurmountable opposition.

Coöperation. A third program of reform, also non-revolutionary in character but much more complex than the two just discussed, remains to be considered. The coöperative movement has taken on such a variety of forms, has spread so widely throughout the world, and has grown to such proportions that it ranks among the most important social movements of our time. It is difficult to frame a definition of coöperation which will be descriptive of all its aspects. The guiding principle underlying the whole movement however is to do away with the profit maker in various departments of economic life. The method is so to organize those

who are subject to the control of the profit maker or who rely upon him for certain services as to enable them to take over control of their own economic activities and perform these services for themselves. A complete study of this great movement would fill many volumes. We shall therefore limit ourselves to a brief summary of the methods and results of three important phases of the movement: producers' coöperation, credit coöperation, and consumers' coöperation.

Producers' coöperation: Description. As developed in connection with our discussion of profit sharing, producers' coöperation is a type of industrial organization in which the workers are at the same time complete or controlling owners of the enterprise. It destroys the wage system entirely by making the workers their own employers and managers. It necessarily destroys the profit income as a separate share of the product by merging profits with wages, and it may cause a similar fusion between wages and the other types of income, *i.e.*, interest and rent. The enterpriser disappears, and his functions are assumed by the wage earners. Producers' coöperation is therefore a type of reform favored by those who condemn the wage system as necessarily unjust to the workers and who despair of any improvement short of complete abolition.

The method of organizing a cooperative productive unit is simple. A number of workers get together some capital, buy or rent some land, and set about producing a commodity under a system of democratic control. Since it can be only by accident that all members of the organization invest exactly the same amount of capital, provision is usually made for the payment of a nominal rate of interest to the workers in their capacity as owners. But ownership carries with it no power. The governing body in the enterprise is elected by the workers, each worker having one vote regardless of his share of ownership or his productive function. By this democratic method managers are appointed and the system becomes one of self-government in industry.

Experience. Enthusiasm for producers' coöperation as a method of social reform has run in cycles during the past century. As early as 1777 an attempt was made in England to establish a business enterprise of this sort. But the first extensive trial of the system

was made in France during the third and fourth decades of the nineteenth century and especially after the revolutionary outburst of 1848. At about the same time the movement made its first real appearance in the United States. In 1833 the cabinet makers of Philadelphia organized a coöperative association; another was formed among the molders of Cincinnati in 1848; and another among the tailors of Boston in the following year. This period also marked the beginning of authentic producers' coöperation in England. The first number of the *Christian Socialist*, published in November, 1850, recorded a needlewomen's association, a printers' association, a bakers' association, two associations of builders, and two of shoemakers in London alone. In no country did one of these early experiments succeed.

In the eighties a new wave of enthusiasm for coöperation arose in the United States during the ascendancy of the Knights of Labor. Many coöperative associations were founded, but again, with one possible exception, they all either failed or survived only at sacrifice of their essential principle. Many self-governing industries were established in England during this period; and at the present time over eighty are in operation in that country. The movement began again in America with the turn of the century, but only twenty enterprises of this character were functioning successfully in 1933. When one examines the remaining examples of successful enterprises in this country and England, he finds that few of them have retained the character of true producers' coöperation, since they allow outside investors a share in the control of management and exclude a part of their wage labor from participation in control.

Cause of failure. The record of producers' coöperation is one of almost continuous failure. The reasons have been much the same in all periods of its history. The wage earners have no adequate appreciation of the difficulties of management and of the rarity of the qualities which make a manager successful. They are loath to pay a wage high enough to attract efficient men to this service. Their own natural leaders leave them because private employers outbid the coöperative association. This same jealousy of the superior man appears in their treatment of the different grades of labor; the scale of wages makes no adequate allowance for the

greater importance of the skilled worker. Above all, democracy has ever shown itself distrustful of concentrated authority and resentful of discipline. Election of managers by those subject to management has always paralyzed the disciplinary mechanism of any organization. These defects, it will be noted, go to the root of the system of producers' coöperation. If the experiments had failed merely through the wage earner's ignorance of the detailed problems of producing and marketing a product, the fault might be remediable. But the weaknesses of this form of industrial organization are essential and fundamental.

Successful experiments which have retained their original purpose are to be found almost exclusively in lines of industry whose operations are simple, where the skill of the workers is fairly uniform, and where serious problems of management do not arise. It may be well to point out that most of the undertakings commonly called "producers' coöperatives" in this country are not examples of true producers' coöperation. Associations of farmers for coöperative marketing or purchasing, coöperative associations of fruit growers, coöperative cotton gins, creameries, and the like are associations of individual profit-seeking enterprises. They may be effective in increasing the profits of their members, but they exert no tendency to modify the wage system.

Credit coöperation. When we leave the subject of producers' coöperation and follow the development of similar movements in other specialized fields of economic activity, we leave behind us the phase of the movement which implies the most far-reaching social reform. Other forms of coöperation appeal to middle class sentiments of thrift and desire for accumulation. They involve little essential change in the economic institutions of modern society, merely an alteration in forms of business organization so that the middleman — the typical profit maker — may be dispensed with. Coöperation in the business of credit and banking is of this conservative nature.

Credit coöperatives, or credit unions as they are commonly called, embody a few very simple principles. They do business only with their own members, membership being easy to acquire by people of good repute through the payment of small dues. Their

funds are obtained from membership dues, from the sale of stock of low par value to the members, and sometimes by accepting savings accounts. These funds are loaned to the small business-men among the members, or to members in need of financial aid for other purposes. The loans require no collateral and are usually at lower rates of interest than are obtainable at the commercial banks. Through these credit unions, small businesses acquire financial resources which would either be entirely unavailable or available only at higher cost. The true credit union is democratically governed; the clients control its policies, and they share in control on the basis of one vote to each person. Any profits from the operations of the union, after paying low dividend rates on the stock, are distributed among the clients in proportion to their borrowings.

The credit union had its origin in Germany in the middle of the nineteenth century. From Germany the system has spread to all other countries of Europe and into the Far East. In 1927 the total number of coöperative credit associations in the world was 85,000, with a total membership of 25,000,000 and an annual business of some $12,000,000,000.

Coöperative credit associations have flourished chiefly in countries with numerous small-scale enterprises in industry and agriculture and with an undeveloped commercial banking structure. But the same principle can be applied to the credit needs of consumers, especially among the wage-earning classes. In the United States credit coöperatives have taken this latter form, their growth having been stimulated by the rapacity of the private loan sharks who exploit the necessitous poor. Their growth in this country has been very slow until recent years. As late as 1921 only ten states had laws permitting their formation; at the present time such laws are to be found in all the states, and the federal government in 1935 made possible their formation under national charters as a phase of the work of the Farm Credit Administration. In 1929 there were about 1,000 credit unions in this country; by 1937 the number had grown to 6,000.

Consumers' coöperation: General description. It is in the field of retail distribution that the coöperative movement has made the

most impressive record of success. Indeed to speak of consumers' coöperation, as it is called, as a mere "success" scarcely does justice to the amazing rapidity with which this type of coöperation has spread over the world and gives but a mild indication of the prodigious scope of its operations. In point of the number of people affected, the scale and variety of its undertakings, and the momentum with which it is progressing, this movement must rank among the most important economic phenomena of all time. Yet its essential principle is simple and its immediate objective modest. Consumers' coöperation aims to displace the retail merchant by transferring his function to the buyers of goods, organized coöperatively. Its object is to save the consumer the profit which the merchant ordinarily receives and thus to permit a slight reduction in his cost of living. The broader undertakings of the consumers' coöperatives have been assumed as an adjunct to this simple function.

The Rochdale system: History. Great Britain is the birthplace of the consumers' coöperative movement and its present leader. From 1830 to 1840 many ventures in this field of coöperation were begun, but all collapsed because of their failure to understand the economic significance of the merchant's function and to hit upon a practicable form of organization. The success of the movement began with the establishment in 1844 of "The Rochdale Society of Equitable Pioneers." This was a group of twenty-eight indigent working men with a vision of great things to be accomplished through unity and coöperation. So great was their poverty that their initial contributions to the capital of the undertaking averaged only a few cents a week apiece. By this laborious process they succeeded, after a year of effort, in gathering together twenty-eight pounds with which to equip a store on a back street in Rochdale. At first they dealt in four staple commodities, flour, butter, sugar, and oatmeal. The store was open for business only on Saturday and Monday evenings, the members serving in turn as salesmen. The following year tea and tobacco were added to their list of commodities, a year later meat. The membership slowly grew until in 1850 it totalled 600. At this point similar associations began to appear in other parts of England, copying the structure and the name of

the original unit. From these small beginnings the movement has spread throughout the entire world.

Principles of organization and management. The Rochdale system embodies certain definite principles of organization and management, among which the following are of chief importance.

(1) *Capital.* The capital is divided into shares of small denomination easily within the reach of the poorest working man. In England the par value of the share is one pound sterling. Each member is required to buy at least one share and permitted to buy more. Shares may be paid for on the instalment plan or even by the automatic accumulation of dividends. A modest rate of return, four or five per cent, is guaranteed on this investment.

(2) *Control.* The policy and management of the association is controlled by the members on the principle of one-man-one-vote. This democratic government prevents a lodging of control in the hands of the larger shareholders.

(3) *Business policy.* Market price is charged for all goods sold, and sales are invariably for cash. Any saving resulting from the superior efficiency of the coöperative principle as compared with private retail merchandising is allowed to accrue, and the proceeds are distributed at the end of the accounting period. Many ventures in this field have attempted to anticipate this saving and reflect it in a cut-rate price policy. These have almost invariably failed. The success of the Rochdale system has been largely due to its conservative businesslike price policy and its avoidance of loss through sales or credit.

(4) *Distribution of profits.* The profits of the coöperative stores are distributed among the members in proportion to their purchases, not their investment of capital. With each purchase the customer is given a receipt, the total of such receipts at the end of the quarter or other accounting period represents his claim to profits. Non-members are encouraged to trade at the stores and receive dividends on their purchases, but usually at a lower rate. As stated above, these dividends may be left on deposit with the store as payments toward the purchase of stock. Thus the system becomes one of automatic saving. The dividends average about ten per cent per annum over a long run period.

These principles are varied in detail in the coöperative systems of different countries, but the basic rules of full market price, dividends on purchases, guaranteed interest on investment, and democratic control are to be found everywhere.

Extent of consumers' coöperation. Only a meagre indication of the spread of consumers' coöperation in England and other countries can be given here. At the present time 1,200 retail establishments are affiliated with the Coöperative Union of Great Britain. The total membership of over 7,000,000 embraces with their families approximately half of the population. The average annual business exceeds one billion dollars. The movement has grown to similar proportions in other countries of Europe and in their colonies and dependencies. In 1914 Professor Gide tabulated over thirty thousand consumers' coöperative societies in Europe alone, with a total membership of ten million people.[1] The movement throughout Europe doubled its membership in every decade until very recent years, when its rate of expansion was checked by the fact that in many countries it already embraced a majority of the total population.

Spread into allied fields. The consumers' coöperative movement of Great Britain has spread out into the fields of wholesale distribution, of production and transportation, of insurance and banking. The Consumers' Wholesale Society of England owns and operates its own steamship lines, has acquired mines, farm lands, tea plantations, and other sources of raw material, and operates factories in the textile, clothing, and food processing trades, which are among the largest in Great Britain. In 1937 it employed 40,000 people and did a business of half a billion dollars. The Banking Department did a total business amounting to over $3,500,000,000 in 1930. The Insurance Department carried over 2,000,000 policies for all sorts of insurance in that year. The Scottish Coöperative Wholesale and the wholesale departments of the consumers' coöperatives of other countries do a similarly diverse business.

Conservatism of consumers' coöperation. How can we account for the amazing success of this great movement? In the first place, it must be recognized that the success of the movement has been

[1] Charles Gide, *Les sociétés co-opératives de consommation*, 1917.

due primarily to its conservative practical policies and not to any far-reaching program of social reform which it contains. The largest single element in its success, at least in England, has been its appeal to the thoroughly conservative desire to accumulate capital and to its provision of an almost automatic method of saving. Retail coöperation has demonstrated beyond all dispute its ability to lower the living costs of the average English family by at least ten per cent. So simple and easy is the process of saving this portion of his normal expenditure and allowing it to grow at compound interest, that the working class member of the system becomes a small capitalist almost without his knowledge.

The conservative nature of consumers' coöperation is further shown in the management of its larger productive undertakings. The system is not one of true producers' coöperation. In its factories and on its farms, plantations, and ships, the Consumers' Wholesale employs ordinary wage labor which is outside its membership. It deals with these workers much as would any other capitalist employer, hiring labor on the basis of efficiency, exercising its right of discharge, and insisting on discipline. It has never adopted a system of profit sharing in these undertakings. Though it has shown more sympathy with the aspirations of the trade union than the ordinary employer and has set the pace in its wage policy, the Wholesale retains in its own hands the rights of managerial control. It has even had strikes in some of its establishments, and it cannot be said to have worked out as yet a policy of industrial relations which is entirely acceptable to organized labor.

Avoiding the wastes of competition. In the second place, the movement has brought to light certain wastes inherent in the prevailing system of marketing and has shown the superiority of the coöperative principle in this field of economic activity. It must be said however that the experience of the coöperators has disproved the common opinion that the ordinary marketing structure is costly because of profiteering practiced by the merchants. The gains of coöperation have come about, not by appropriating an excessive private profit, but by reducing costs inherent in competitive enterprise. The marketing side of our productive system has responded less than any other to competitive pressure

toward greater efficiency; herein lies the opportunity for coöperation. Since the customers of the coöperative store control its buying, they know in advance what kinds of goods will be in demand and in what quantities, and they can adopt the least expensive means of getting these things, avoiding also the loss of unsold stock. The shops are sure of their clients' patronage and do not have to practice expensive policies of advertising and salesmanship to attract custom from other merchants. Advertising and selling devices, when used by consumers to facilitate the satisfaction of their wants, have a different function from that of a system of marketing where competing producers contrive to make the consumer want what they have for sale.

The wastefulness of competitive marketing is especially great in the older countries of Europe, where the social stigma attaching to this form of activity tends to repel the able enterprisers and leave the function in the hands of mediocre men. In these countries the customary practice of extending credit to customers has been a source of loss; business methods in retail merchandising have been traditional, lethargic, time-consuming; shops have been managed by rule of thumb. Accordingly the opportunity of improvement through a system managed with greater alertness and on stricter business principles is especially great. Consumers' coöperation has attained its maximum growth in such countries.

Broader social aspects. What has been said about the conservative and narrowly practical character of consumers' cooperation must not be taken to mean that this movement is not a power among the broad social forces of the day. The movement is distinctively working class in its origin and present personnel. It is the largest single working class organization in many of the countries in which it has gained a foothold. Of necessity, its members develop an *esprit de corps* which can be utilized to unite them for collective action along other lines than that of mere retail merchandising. Because of its democratic control, the coöperative movement is especially susceptible to the idealism and the reform tendencies of the rank and file of humanity. Hence we find it involved in many political and social activities which are quite apart from its strict function. In these phases of their policy the coöperatives

take their tone from the prevailing temper of their membership. They will be socialistic in countries where the mass of working people are radical, mildly progressive in more conservative countries. These broader activities do not spring from the principle of consumers' coöperation, but merely represent the universal tendency of social institutions to take on new functions as they are made the medium for expressing the many sided interests of their members. Yet so gigantic has this movement become that these supplementary activities must be reckoned among the most powerful of modern social forces.

Social program in England. The coöperatives were leaders in working class education in England. From the beginning they appropriated two and a half per cent of their profits for educational purposes. They now hold scholarships at Oxford University, support liberally extension schools and colleges for working men, and have recently established a college of their own at Manchester. In the conflict over industrial relations they are on the side of labor. They have loaned funds or supplied food to men on strike, and their aid was a large factor in promoting the solidarity of labor and increasing its strength during the trying period of readjustment which followed the World War. Politically the Coöperative Union supports the Labor Party, with its program of constitutional socialism. Its influence on radical movements in England was shown in connection with the development of the Building Guilds. These organizations, producers' coöperatives with broad socialistic objectives, were saved from collapse at the outset through the willingness of the Banking and Insurance Departments of the Coöperative Union to finance and underwrite their enterprises.

In other European countries. The coöperative movements of the other European nations have been swept by the idealism of their members into similarly far-reaching social activities. In Belgium the profits of the coöperatives are not distributed but are invested in the Maisons du Peuple. These are great community centres, supplying the working classes with facilities for amusement, education, and cultural activity. The Maisons give free medical and clinical service, provide relief for men on strike, organize and finance travel parties for young people at home and abroad.

The movement in France has strong political affiliations and is recognized in the administrative machinery of government through a Council of Coöperation in the Ministry of Labor. So it is with other countries. The forms taken by these social and political phases of the movement as well as their effects on the established institutions of capitalistic society vary in conformity with the peculiar social conditions of the different countries and with the prevailing temper of the working classes. But everywhere the movement has become a power among the social forces of our time.

Obstacles to consumers' coöperation in the United States. America has remained largely untouched by the movement whose impressive record in other lands has been sketched in the preceding pages. Until 1900 most American ventures in this field were short-lived; nothing approaching a movement of national scope existed. From 1900 to 1914 the appearance of coöperative retail stores clustering about the chief industrial cities gave the first real evidence of vitality in the movement. During and after the World War retail coöperation began to spread with some rapidity, until the development was checked by the depression of 1921. Since that year there has been a slow but steady growth of the movement in this country. At the present time there are in existence some 11,000 retail coöperatives, some 50 societies for coöperative construction of houses, and about 40 wholesale societies serving these different ventures. The total membership is nearly three million, and the total volume of business approaches half a billion dollars annually. These are large figures, but they really show a small record of achievement in comparison with the situation in other countries. The coöperative movement in this country has not branched out into the productive lines in which that of other countries is so prominent. It is noteworthy also that the movement flourishes principally among foreign born wage earners who have been trained to the system in other countries.

The reasons for this slow growth in the United States are inherent in conditions peculiar to this country. They may be summarized as follows: (1) Relative efficiency of the marketing system. Retail merchandising has been more progressive than in Europe and has attracted many of the ablest enterprisers. Department

stores, chain stores, and mail order houses leave little room for saving of cost in distribution. The coöperatives of England are just now beginning to feel and to fear the competition of these forms of efficient profit-making enterprises. (2) The great extent of the country has been a handicap to any unified movement among consumers. (3) The greatest obstacle has been the individualistic and impatient character of the American population. There has been no stable wage-earning class with sufficient solidarity to support such a movement. The tendency to migration has continually broken up local community groups. The higher standard of living and the greater ease of life have prevented the development of the thrifty, penny-saving habits to which such a movement as this makes a strong appeal.

LIV

SOCIALISM

Radical movements. The reform movements described in the preceding chapter would leave unaltered the essentials of the capitalist system — private property, profit-seeking enterprise, competition — and seek to modify their effects to the betterment of certain groups in society. There are other organized movements far more revolutionary than these, which aim at a complete reorganization of the system itself. These revolutionary programs are of great variety, divided frequently within themselves into factional groups, which conflict with one another over matters of doctrine and tactics. It will suffice for our purposes to group them all under the general title, *socialism*, reserving for special consideration one large section of the revolutionary movement, *communism*, which, though originating in the same body of doctrine, is now operating independently of the other socialist groups. The Fascist program will also be considered in this connection, because of its close association with revolutionary socialism, although it is commonly — and erroneously — classified among the conservative movements of our time.

Non-economic features of socialism. Socialism is a complex and many-sided movement, including so great a variety of practical and idealistic aspirations and assuming such diverse forms in different places and at different times that it is difficult to define it in a single formula and quite impossible to describe and analyze it fully in the small space at our disposal. It is to be fully understood only as a strong current in human thought and behavior, involving millions of people and carrying them along different paths of social change. Its rational program is not its most important feature. Its emotional content, the idealism of its followers, its dynamic power, are the features of chief significance from the standpoint of its present and future influence upon the world's social organization. Since these features cannot be brought to light by a critical

analysis of the economic doctrines and the paper programs of socialist parties, they receive too little attention in most studies of this movement. We too are obliged to leave them out of account in the discussion which follows. But it is well to begin our study fully aware that no adequate appraisal of this movement can be based on economic analysis alone.

Definition of socialism. Socialist doctrines contain four groups of ideas: an indictment, an analysis, a program of reform, a campaign. The *indictment* is a list of charges against "any and all industrial systems based on private property and competition." The *analysis* traces the development of capitalism back to its origin, dissects its present structure, and stresses certain developmental tendencies which, so it is argued, must inevitably destroy our present order. The *program* outlines a substitute for the capitalistic system which socialists believe will remove the evils of modern capitalism. The *campaign* is a plan of immediate tactics for those who adhere to its cause.

Agreement among all socialist groups is not to be found in any one of these sets of ideas, though there is greater unanimity of opinion regarding the indictment of the capitalist order and the analysis of its structure and development than in the details of the program or the plan of campaign. In their indictment of capitalism socialists differ only in the breadth of their condemnation of the present order and the relative stress they place on its different faults. Since the socialist analysis of our institutions is derived almost entirely from Karl Marx's classic interpretation of capitalism, there is considerable agreement here.

Socialist programs differ widely in detail, but we can discern an essential similarity among them. They all propose a type of society in which there will be no private property in wealth used for productive purposes, and they agree further that the productive system of the future shall not be governed by competition and private initiative, but by some sort of collective control and guidance. Socialism then is destructive of these two basic institutions of our order, private property and free competition or individual enterprise. The programs differ in their substitutes for these institutions. What group shall own the productive wealth of

society when private ownership has been forbidden? And through what mechanism shall society decide upon the kinds and quantities of goods to be produced and fix their prices, when these things are no longer left to the free play of competition? It is over these questions that the socialist movement breaks up into different schools.

In regard to their campaign socialists divide into two groups: those relying on peaceful, evolutionary methods, and those who place their reliance on violence and force. The former are *constitutionalists,* believing that gradual increase in their power on the political field will eventually bring them to their goal by thoroughly legal means. The latter are *direct actionists;* they have no faith in politics, being convinced that only through force or threat of force will the revolution come. The official socialist political parties in most modern democratic countries belong to the constitutional wing of the movement. Examples are the Socialist Party in the United States, the Labor Party in England, the Social-Democratic parties in the Scandinavian and other European countries. The outstanding example of the direct actionists is the Communist Party in Russia and in all other countries where this party exists.

Any definition which attempted to embrace all these phases of the socialist movement would be too cumbersome for practical use. We shall define the movement in terms of its program, stressing the essential similarities between the different socialist schools. *Socialism is a program of reform which repudiates private ownership of the means of production and competitive control of industry; it proposes a type of society in which productive capital will be owned collectively and economic activity will be controlled by authority.*

Socialist indictment of capitalism. The socialists have no quarrel with capital. They believe that the supply of capital will be increased and its efficiency improved under their system. Their attack is upon capitalism — the organization of society — not upon the material instruments of production. Like all reform movements, socialism has its "lunatic fringe," whose indiscriminate denunciation of all that is and whose panegyric of the Utopia of the future do not merit serious consideration. There are however many points in the socialist indictment which are thought-provok-

ing to every student of society; these are summarized in the following paragraphs.

(1) The competitive system is wasteful of resources and effort. In established concerns individual enterprise results in duplication of effort. Competitive advertising, competitive salesmanship, and other forms of competitive activity increase the costs of each enterprise. These costs could be reduced if the services common to all producers were provided through a common agency. In the competitive system the superior business methods and most efficient machines are controlled by individual competitors for private profit. Increased efficiency for society as a whole would result if the best processes were made the common possession of the entire industry. Thus far the argument is much the same as that which contends for the superior efficiency of combination and large-scale production. In a more general sense the socialists urge that the present system is wasteful because enterprisers, in their search for profit, direct capital and labor into fields of activity which are already well enough equipped to serve the needs of society. It would be more economical to use collective intelligence in forecasting our needs and conscious coöperation in adjusting the means of production to the ends in view.

(2) The competitive system results in wasteful strife. Employers and workers strive with each other over the division of the income from industry. Producer strives with producer for advantage in the market. Out of this strife spring practices which result in loss of productive forces, such as were disclosed by our study of industrial relations and of governmental control of competitive practices. The anomaly of the situation is that these contesting parties are really coöperating toward common ends; that is, they are all laboring to satisfy the wants of mankind. Intentional coöperation is less wasteful than antagonistic coöperation. From a broader viewpoint, socialists believe that the competitive system is the root cause of international war, with its incalculable toll of destruction and waste. This is because, in their opinion, it is the competitive struggle for profitable foreign markets and for sources of raw materials which leads governments to fight each other in behalf of the business interests of their people. They

believe that their system will automatically produce harmony between individuals, social classes, and nations.

(3) The competitive system is destructive of human and natural resources. This criticism is implied in part in the foregoing statement. But in a more general sense it is urged that the pursuit of immediate private gain leads to the exhaustion of irreplaceable natural resources and the wasteful exploitation of human energy, especially that of children and the more impoverished adult workers. The modern conservation movement and modern labor legislation imply a similar criticism of uncontrolled competition.

(4) The competitive system is unstable and erratic. Socialists argue that the business cycle is the result of distorted production, which in turn results from anarchy in production and inequality in distribution. Seasonal and cyclical fluctuations, with their recurrent losses from idle capital and labor, could be prevented in a system controlled and guided collectively. Less comprehensive maladjustments which occur from day to day could also be prevented through the application of intelligence. Thus there need be no sudden gluts in certain markets, with their consequent loss of values, and no unexpected dearths, with their attendant deprivation of human satisfaction, if production and marketing were carried on according to plan.

(5) Private property results in great inequality. Private property includes the right of inheritance. As the socialists view the matter, the great inequalities in incomes are caused by private ownership of income-bearing wealth and perpetuated by the institution of inheritance. The socialists believe that capacity is distributed with approximate equality among the different income classes. Opportunities for self-development cannot be equalized so long as some men are born to a position of privilege while the many are born in an environment of poverty, ignorance, and hopelessness. This situation is not only unjust on ethical grounds but is inefficient and precarious; inefficient, because it smothers the latent abilities of those in the lower groups and subsidizes idleness in the upper classes; precarious, because it arouses the envy and resentment of the victims of the injustice, who are an overwhelming majority and have the power to destroy the system.

(6) The wrong things are produced. Inequality of distribution distorts the whole productive system out of alignment with the real needs of society. Only those things are produced which will sell at a profit. Hence the kinds and quantities of goods resulting from the combined activity of mankind are determined by the spending power of individuals and not by the test of what is needed to nourish a strong, healthy population. There are abundant examples of the grotesque effect of mal-distribution on production. A serious housing shortage may exist at a time when much labor and capital are being devoted to the construction of steam yachts and country palaces. The average family may suffer from a shortage of nourishing food and serviceable clothing while useless or injurious articles of diet and frivolous articles of fashion are supplied in abundance to the rich. Foolish expenditure of all classes has the same distorting effect on production. The production of chewing gum, cosmetics, and patent medicines absorbs capital and labor which ought to be supplying impoverished people with the means of wholesome living. The same evil is to be seen in the fact that many things of vital importance to human welfare are not produced at all, because no individual enterpriser can make a profit out of their production. The socialist system proposes a quite different basis of production. The social group will decide what things are needed to support a wholesome life, and these things will be produced first in such abundance that there is plenty for all. Only surplus labor and capital will be devoted to the production of luxuries and extravagances, while injurious goods will not be produced at all.

(7) Capitalism is destructive of liberty. The defenders of the present system of industry are accustomed to speak of it as a system founded on individual liberty. It is said by the spokesmen for capitalism that our society permits each human being to work out his destiny according to the powers that lie in him unrestrained by bonds of status and tradition. This the socialists deny. To them the present order means the enslavement of the great majority of propertyless people by the owners of wealth and the subjection of the whole body of society to the dictates of those who dominate its essential industries.

First the wage system is condemned as a species of slavery. It is explained as an arrangement between the few who own the instruments of production and the resources of the earth and the many whose only hope of survival consists in their getting access to these things. The outcome of the arrangement is that the many deliver themselves into the power of the few, their productive energies to be exploited to the profit of the masters. The freedom of the individual worker is freedom to starve if he does not wish to surrender. It is not charged that the employers necessarily wish it to work out this way or that they are even conscious of the result. They too are caught up in a system which inevitably produces the subservience of the "have-nots" to the "haves."

Secondly the economic destiny of the social group as a whole is at the mercy of the more ruthless or more fortunate individuals. The landlord exacts his tribute for the privilege of using the gifts of nature. Those who have acquired or inherited ownership of socially necessary instruments — railroads, steel mills, factories — can exert a large measure of control over the rest of mankind. These essential instruments are social in their purpose and importance; they cannot safely be allowed to fall into the hands of individuals.

Finally all social classes — wage earners, employers, capitalists, landlords, consumers — are seen to be slaves to the passions of self-interest and acquisitiveness which the competitive system arouses. The socialists believe that the order of society which they propose will be one of real freedom. Men will not work for the profit of other men but for the group as a whole and under its commands. The instruments of labor and the resources of nature will be owned and managed collectively under democratic government. Equality of opportunity and free choice of self-development will not be denied any individual because of mal-distribution of wealth. The passions of greed and self-seeking will be held in check by arrangements which play upon better qualities in human nature.

Socialist Analysis of Capitalism: Karl Marx. Turning now from their indictment of the evils of capitalism to the socialists' analysis of the development and present tendencies of modern society, we

find this phase of the movement to derive in the main from the work of Karl Marx, who was the founder of almost all the modern schools of socialism. The *Communist Manifesto,* written in collaboration with Frederick Engels in 1848, and *Capital,* an analysis of the economic laws of capitalism and a forecast of its future development, which appeared in 1867, supplied the doctrinal basis for all subsequent radical movements. Later expositors of socialist ideas, such as Kautsky of the Social Democrats and Lenin of the Communists, have almost without exception sworn allegiance to the Marxian doctrines. It will serve the purposes of our study therefore to summarize here the basic theories of Karl Marx. These theories may be classified under three heads: his materialistic interpretation of history, his system of economics, and the laws of capitalist development which he outlined.

Materialistic interpretation of history. Karl Marx was a forerunner of the modern group of historians and sociologists who explain societal evolution in terms of economic factors. Throughout all human history the framework of social institutions, according to this doctrine, has derived its peculiar character from the way in which the people have made their living. Government, religion, family system, art, literature, and morals are by-products of man's economic activities, harmonizing with the prevailing system of production and exchange. The major changes which have appeared in these institutions in the course of societal evolution have been the effects of prior changes in the economic basis of society. The most important events in human history — wars, the rise and fall of empires, the Crusades, the discoveries, and the migrations of peoples — are traceable to economic causes. This doctrine, in some degree of comprehensiveness, in now accepted by most students.

The class struggle. But Marx carried the doctrine one step further. To him not only were the economic institutions basic and determining, but these institutions had their own vital centre in the class struggle. Every system of production which mankind has hitherto devised has created a hierarchy of classes, exploiters and exploited. In classical society slave, plebeian, patrician, and noble; in the Middle Ages serf, vassal, knight, and feudal lord,

were ranged in an ascending order of increasing privilege. Not only is class antagonism the essential characteristic of any historical social order at any given time, but this antagonism has produced the dynamic power which has caused society to evolve and change its structure. For the interests of the classes have clashed and forced them into continuous conflict. Revolutions in economic institutions, themselves the cause of all great changes in other institutions, have resulted from the class conflict when a previously exploited class gained power to throw off the yoke of the oppressor. Thus "the history of all hitherto existing society is the history of class struggle." But each such revolution merely reëstablished the system of exploitation in a new form; for the victorious class in turn organized a system of production with itself in the seat of power and the rest of mankind subservient to it.

The last great change of this type occurred when the bourgeoisie arose and overthrew their feudal masters, thus instituting the system of capitalism. Out of this grew the present class alignment, capitalists and laborers, and the present struggle between these two classes. The outcome of this struggle lies in the future, but it is predicted that the wage earners will eventually succeed, as did the present capitalists, in overthrowing their exploiters. Their triumph will mark the end of the class struggle, since they will reorganize the industrial system so as to remove the possibility of exploitation. This will be true because the wage earners, being a vast majority of the whole population and committed to coöperative action, will organize the economic structure on collectivist lines, in which each will work for the welfare of the whole group. Thus the socialist revolution will usher in a classless and therefore stable society.

Marxian economics. The base of Marx's system of economics is the labor theory of value. Normally and in the long run, he says, goods are exchanged for each other in ratios which measure the relative amounts of labor power incorporated in them. Various exceptions are acknowledged, and it is admitted that market prices may temporarily depart from this value norm. But on the whole it is labor cost, including in this concept a proportionate part of the labor cost of providing capital instruments and raw materials used in production, which controls value.

The law of wages is a corollary of this law of value. Labor, like other goods, sells at a price equal to its cost of production. The wage rate is determined by the cost of "the necessaries required for the laborer's maintenance plus another amount required to bring up the quota of children who are to perpetuate the race of laborers." Allowance is made for the standard of living, but only in so far as it affects the birth rate. This is the "iron law," or cost of subsistence theory, of wages.

These doctrines of value and wages, according to Marx, give the clue to the process of exploitation in capitalistic society. The wage earner is paid on a cost of production basis, but the wage contract transfers his entire productive power to the capitalist. He is able to produce enough to replace his own subsistence, and therefore his wages cost, by working a few hours a day; but being at the mercy of the owner of capital, he is compelled to work for a longer period, his surplus products being appropriated by the employer. This part of the product therefore is obtained by the capitalist without cost to himself. It is sold, like other products, at a value determined by the amount of labor it contains Here we have the concept of *surplus value*, which in the Marxian analysis epitomizes exploitation. The incomes of interest, rent, and profits are drawn from the sale of this surplus product of the wage slave. The recipients of these incomes are parasites, exploiting, not the consumers as is sometimes believed, but the laboring class.

That this whole structure of theory is unsound will be apparent to every student of economics. The value theory is fallacious, because it leaves out of account the influence of marginal utility. In a competitive system such as Marx was describing wages tend toward equality with the marginal productivity of labor and are not determined by subsistence costs. This explanation of wages disposes at once of the vitally important concept of surplus value. Opinions may differ as to whether interest, profits, and rent are justifiable incomes, but in any case the Marxian explanation of their existence and their amount is erroneous. Until very recently socialists have been content to repeat these economic fallacies; now they are beginning to perceive their inaccuracy and to discard

them. The strength of modern socialism lies elsewhere than in its system of economic theory.

Marx and the Utopians. The conclusion of Marx's diagnosis of the capitalist order was a conviction that socialism is the inevitable next step in societal evolution. His system made a sharp break with all preceding socialistic programs. Before his day socialism had taken the form of Utopia building, the construction of little communities in which a perfect order of society was to demonstrate its power to evangelize the world. The Utopians believed that all that was necessary to usher in socialism was suddenly to call a halt on the operation of the present system and by persuasion to transform it over night into an ideal commonwealth. Their experiments failed, and their propaganda lost its appeal to all but the most visionary. Marx repudiated with scorn this whole system of thought. It is quite as impossible, he said, for the prophet and the evangelist to produce socialism artificially as it is for the alchemist to turn baser metals into gold. Socialism must come as a natural outgrowth of capitalist society, or it will not come at all. Socialists may speed up the intermediate stages leading to the new society and may facilitate the transition, but the motive power must be supplied by the forces of capitalism itself.

Laws of capitalist development. Marx thought that he had discovered the existence of two such motive forces actively tending toward socialism and a third characteristic of capitalism facilitating the eventual transition to socialism. The active forces were: (1) the tendency toward concentration and (2) the increasing number and misery of the laboring class; the facilitating factor in capitalism was (3) the recurrent collapse of the productive system in times of crisis.

Tendency toward concentration. As to the first of these forces, it is argued that capitalist development tends in two directions toward increasing concentration: first the concentration of industry into larger and larger business units, secondly the concentration of wealth and income in the hands of a diminishing number of people. Industrial concentration is said to be due to the superior efficiency of large industries, which enables them to undersell and crush out

smaller competitors. We must note to Marx's credit that this prediction was made before the trend toward large-scale production had set in. But Marx set no limit to the process, leaving his readers to infer that concentration would continue until a single gigantic business unit monopolized each of the major fields of industry, including agriculture. In agriculture the prediction has been falsified by subsequent events. How far it has been verified in other fields we have seen in another place.[1] This tendency was to lead to socialism for the double reason that collective control would become both necessary and easy; necessary in order that society might protect itself from the menace of these monopolies; easy to establish because of the disappearance of effective competition, and easy to administer because of the relative simplicity of operating the few centralized productive units.

Concentration of wealth was inferred, partly from the preceding tendency, partly from other evidence. Marx made the mistake of assuming that concentration in industry of necessity involved concentration of ownership. He wrote before the day of the corporation, whose joint stock principle permits a growth in the scale of the business unit coincident with an increase in the number of those who hold property rights in it. Great as are existing inequalities in distribution, we have no evidence that they have grown greater since Marx's day or are now growing greater.

Increasing number and misery of the laboring class. The tendency toward concentration, coupled with other features of the capitalist system, was expected to produce an expanding propertyless class, with a steadily declining economic status. The crushing of small enterprises by the growing monopolies would precipitate former small capitalists into the laboring class. The growth of an industrial reserve army, the effects of periodic depression, and the absorption of an increasing fraction of income by the rich would cause increasing misery for this ever-growing group. The rich would grow richer and fewer, the poor more numerous and poorer. This process would cause the system to break down through sheer inability of the masters "to maintain their slaves in their slavery." It would also cause the victory of the proletariat to put an end to

[1] See especially Chapter IV.

the class struggle, since it would be a victory of "the immense majority in the interests of the immense majority."

It is evident that the part of this forecast concerned with the mere size of the wage-earning class proved true subsequent to Marx's day. During the following decades the laboring population in modern industrial countries grew both in absolute numbers and in proportion to the whole population. But the more important part of the prediction — that of increasing misery — has not been verified by the event.

Marx did not foresee the growth of trade unionism, social insurance, and labor legislation, which would protect the most defenseless of society's members from the consequences of their own bargaining weakness. At least until the opening of the twentieth century the average workman's absolute income increased steadily, and his relative position in distribution at least did not decline. The widespread unemployment and destitution which have appeared in the past two decades might indeed lend color to the belief that the condition of the masses at the present moment is one of increasing impoverishment. But this period has been too disturbed by war and by business depression and has moreover been of too brief duration to serve as the basis of any long-range prediction.

The rôle of crises. Marx made much of the tendency of industrial society to fall periodically into crises and depressions. In his opinion crises were the inevitable outcome of the exploitation inherent in the capitalist system. The wage system resulted in withholding from industrial laborers a substantial fraction of the values attributable to their productive efforts. This would be momentarily profitable for the employer, but for the system as a whole it would mean that goods were continuously produced in excess of the ability of the great mass of the people to take them off the market. This situation would run along for a time until the accumulated stores of unsalable goods precipitated a crash, followed by a depression, during which production would stop until the surplus of products had been liquidated. Marx believed that such paroxysms of the industrial system would not only recur periodically but would grow progressively more severe.

This theory of crises held a place in Marx's prediction of the inevitability of socialism. In the first place, each crisis and depression would play its part in increasing the number and the misery of the propertyless classes, thus promoting the trend described in the preceding paragraphs. Moreover as each collapse became more complete and recovery more delayed and difficult, the time would come when the recuperative powers of the system would prove inadequate to the task of restoring equilibrium. The crisis itself was a challenge to the validity of the capitalist system, calling attention to faults which permitted an excess of the products of man's labor to exist coincidentally with widespread misery and want among the laborers. It therefore predisposed ever-increasing numbers of people to turn against the system and seek for a substitute. There would come at last a collapse so cataclysmic that revolt would occur before the protagonists of capitalism could bring about recovery. Both Marx and his immediate followers believed that the transition to socialism would come about during a business crisis.

No dogmatic judgment can be passed on this theory of crises. The business cycle is a baffling feature of the modern industrial order, concerning whose causes and remedies there is no agreement among those who have studied the problem exhaustively. None, it may be said, believes with Marx that the essence of the situation is the existence of surplus value as defined in the Marxian doctrine; indeed the socialist scholars themselves no longer hold to this doctrine in its original form. Marx's theory of the cause of crises and depressions will however be recognized as a variant of the theory of overproduction, to which some experts who are not socialists subscribe in one form or another. There is a respectable body of argument to the effect that industrial disequilibrium, in the absence of disturbing factors such as war, is due to the failure of the productive system to return wage incomes large enough to enable the great mass of the people to buy the product of industry at profitable prices. Other students of equal authority deny the validity of this explanation of the business cycle.

But the more important part of Marx's theory, from the standpoint of predicting the inevitability of socialism, is its forecast of

the increasing severity of crises, ending in the final destruction of the capitalist system. This has not been confirmed by subsequent history. Crises and depressions have varied in severity and duration, with no pattern either of progressive increase or of progressive decrease discernible as a trend. It is true that the present business depression has been of longer duration and more comprehensive than most previous ones; and there are those who undertake to prophesy that society will not recover without fundamental changes in its institutions. But experience does not support such prophesy; the social order has heretofore survived these periods of painful readjustment without essential changes. Periods of distress, by increasing discontent, do strengthen the socialist movement temporarily, but in the past this influence has been offset by the return of better times. It is probably true that the capitalist order forestalls popular revolt at such times as these only by its willingness and ability to relieve distress at public cost, thus preventing hardship from developing to the point of desperation. But this really argues the vitality and adaptability of capitalist institutions, not their impermanence. It is noteworthy that in Russia, the one country where successful socialist revolt has occurred, the ground was laid for it, not by business crisis as the term was used by Marx, but by the chaos and confusion resulting from a disastrous war.

General appraisal of Marxian doctrine. Although, as has been pointed out in passing, there are serious errors in Marx's analysis of capitalism and many of its specific predictions must be dismissed as unproved, there remains much of value to the student of social evolution. Marx's basic interpretation of history has provided a convenient and illuminating approach to an understanding of the historical process. His analysis and forecast of capitalist development, written as it was in the early stages of industrial society, displayed rare powers of insight.

But the Marxian doctrine is important chiefly for its influence in shaping the development and the present character of socialist thought. It placed socialism on a new and firmer basis. There is real validity in the central thesis; *i.e.*, that every thorough evolution in society's institutions must come about as a natural unfolding

of forces at work in the preceding order. Until the rise of communism this principle had great influence on the practical tactics of all the followers of Marx. It convinced them that capitalism must run its course before socialism is possible. The factory system in industry, with its corollary of a large wage-earning class; democracy in the political field; education for the mass of mankind; experience in collective action gained through labor organizations — these phases of capitalism must be nourished to maturity. Violent premature revolutions would fail disastrously. Marxian socialists were forced by the logic of their system of thought into peaceful and legal forms of activity. Many of them believed, as did Marx himself, that the final overthrow of capitalism could not occur without violence and disorder; but their preliminary program was evolutionary rather than violent.

This outlook continues to dominate the doctrines and tactics of all the principal socialist groups except the communists. These revolutionaries, although they claim Marx as their spiritual leader and profess to derive their doctrines from him, are convinced that socialism can be established even in backward countries which have had no experience of capitalism. They also deride the peaceful political tactics of other groups and place reliance on direct action. So complete is this break in doctrine and strategy that it would be misleading to attach the term, socialism, to both wings of the movement. Accordingly we shall deal with them separately in the discussion of program and campaign which follows.

The socialist program. We have seen that all socialist groups agree in that they propose to abolish private ownership in the means of production and to substitute for competition some form of authoritarian control over productive processes, but they differ as regards the form of collective ownership and control. The principal line of division is between (1) those who would make the present territorial state the owner of productive instruments and would vest control of economic activity in the chosen representatives of all the people and (2) those who would place each branch of industry under the government of the workers in the industry, relying upon various devices of federation to attain harmony within the economic system as a whole. The syndicalists and guild socialists are examples

of the latter branch of the socialist movement. Their ideas have attained some prominence in a few countries; but on the whole they are far less influential than are the representatives of the other wing of the movement, who take the name state socialists from the pivotal feature of their reform program.

State socialism does not demand complete centralization of society's economic life in a single unit of government. Its program provides for division of ownership between different political units. The central government would own the industries whose significance is nationwide; to the smaller territorial divisions and the municipalities would belong those which supplied only the local markets. Some allowance is made for individual enterprises, provided they refrain from employing hired labor and sell their products under strict price control. In this system, when fully established, wages would be the only possible type of income, and everyone who wished to escape starvation would be compelled to work, excepting of course the aged, the children, and the defectives, who would be cared for by the state. The state would be the dominant power in the lives of all; it would be the only employer, the owner of all land, the only saver of capital, the only seller of goods. It is intended that this great power shall be brought completely under democratic control through universal suffrage, direct election for short terms of all policy-making officials, proportional representation, single chamber legislatures, and other devices.

State socialism includes those who accept the central principle of political ownership and control but modify in some details its more extreme doctrines. In fact the majority of state socialists are not doctrinaire but opportunistic in outlook. They press for an extension of government ownership and control, wherever private ownership appears to them to threaten popular welfare and competition seems ineffective or destructive, without much concern as to the final outcome of the process. They are willing to leave as large a sphere for individual initiative as is consistent with the general principle that land and the larger aggregates of capital be used for the "common good" rather than for private profit. The Fabian Society of England is best representative of this school of

modified socialism. This small group of intellectuals, which has numbered among its members such people of international prominence as Sidney and Beatrice Webb, Philip Snowden, Bernard Shaw, Graham Wallas, H. G. Wells, and Ramsay MacDonald, has given much study to practical administrative problems of socialism. They have supplied the intellectual leadership and much of the program of the British Labor Party and have provided members of the Labor cabinet during that party's tenure of power.

In most democratic countries socialism as a practical political movement has taken on this moderate, temporizing character. Political parties are formed locally and nationally to agitate for an extension of government ownership and regulation of industry. The immediate objectives are usually limited: municipal ownership of the principal services; nationalization of railroads, mines, and public utilities; legislation to limit the power of employers and strengthen organized labor, and the like. The peaceful methods employed and the relatively minor changes proposed at any one time often conceal the full implications of the movement. But this gradual encroachment on the domain of capitalism is expected eventually to produce an economic order so extensively socialized that the establishment of a thorough-going socialist state will be easy and natural, and it is in terms of these final results that the whole movement must be judged.

The socialist campaign: Organization and policies. The modern socialist movement may be said to have been launched in 1864, when Karl Marx founded the International Workingmen's Association to bring together into one organization the social revolutionaries of Europe. The membership was diversified, containing anarchists, political reformers, and others with vague and ill-formed programs, as well as the small group which understood and subscribed to the doctrines of Marx. Conflicts of policies and the jealousies of rival leaders prevented this first international organization from attaining internal unity; and the disastrous outcome of the Paris Commune, in which the Association was represented, finally caused its disruption. During the two decades following 1871 the international socialist movement was divided into rival camps, but the ideas and policies of Marx were steadily making

headway, and in 1891 the two leading factions of the movement coalesced to form the Second International. This organization consisted of two delegates from each constituent national party. The local parties remained autonomous and, as they are in the socialist movement today, free to adopt diverse policies and tactics in harmony with national conditions. But the central organization provided a medium through which they could reach agreement on broad questions of doctrine. Under the leadership of the Second International the socialist movement in most countries turned to methods of peaceful political action, eschewing the violent tactics of the First International.

The Second International was destroyed by the World War. It had adopted resolutions in 1910 committing its constituent national parties to opposition to their governments in case of war and to coöperation among themselves during any conflict. When war came patriotism proved too strong for these international aspirations in the two key countries, France and Germany, the parties in these countries voted to support their governments, and the International disappeared. Abortive attempts to revive it during the war failed, and it was not until 1923 that international union was finally reëstablished under the name The Labor and Socialist International. In the meantime Lenin had formed at Moscow in 1919 the Third International, discussed later, to unite the communist parties of the world. The Labor and Socialist International remains the central representative of conservative socialism throughout the world. Its headquarters are at Zurich, Switzerland. It is with this organization that the Socialist Party of the United States is affiliated.

The development of socialism in the separate countries is too complex a subject to be given adequate treatment here. Everywhere the movement is divided within itself over questions of doctrine and tactics into factions, which in most countries are organized for political purposes in opposition to one another; this quite apart from the separate organization of the communists, who are bitter antagonists of the socialist movement as a whole. The divisions within the socialist movement proper usually reflect conflicts over the relative emphasis to be placed on organized labor, on the one hand, and the territorial representative state, on the other, as

the sovereign body in the socialist economy, or differences of opinion as to the rapidity with which modern industrial society should be socialized. Without going into detail it can be said that the Social Democratic Party, to be found in many countries, the Labour Party of England, and the Socialist Party of the United States, usually representing the majority wing of the movement, stand for reliance on political government and a cautious experimental rate of change; while the syndicalists and the radical socialists favor revolutionary action through labor unions and a hasty continuous drive toward socialization.

Political action. The dictatorship countries, whether communist or Fascist, have crushed out a once strong socialist movement within their borders. In other industrial countries, excepting the United States, socialism is a powerful factor in the political and social life. The socialists have controlled the governments of many municipalities throughout Europe and in Great Britain. They always form a strong opposition party in the national assemblies and at times have controlled the central governments in European democratic states. Typical of conservative socialism is the Labour Party of Great Britain, formed early in this century as the political arm of British organized labor. From small beginnings this party grew in twenty years to become the strongest single political group in the country. For a short time in 1924 and again in 1928, though still a minority party both in terms of popular vote and representation in Parliament, the party was called upon to form the ministry and conduct the affairs of the nation.

In the United States socialism has never been strong as a political movement. The individualism of the people, fostered by the relative abundance of economic opportunity in a new and naturally rich country, the retarded development of organized labor, and the peculiar constitutional structure of the national government, which places almost insuperable obstacles in the path of minority third parties, are among the factors accounting for this situation. In local elections the socialists have at times gained control of municipal governments, notably in Berkeley, Bridgeport, Butte, Milwaukee, Reading, and Schenectady. They have sent representatives to many state legislatures and have elected members to the

House of Representatives. But in presidential elections, in which they have engaged since 1904, their record popular vote, attained in 1920, was less than a million, and it has declined since then. Nevertheless the influence of socialist ideas upon governmental policy in this country has been greater than is indicated by the political record of the party. Many reforms — such as social legislation, legal immunities for trade unions, governmental regulation of business, and direct government activity in economic enterprise — originally sponsored by socialists, have come to be adopted by the major parties and put into effect in the several states and in the nation. There is in fact little difference to be found today between certain policies of the more progressive of the major parties and the immediate program of the Marxian socialists advocated by them as steps leading gradually to complete social reconstruction.

Communism. Before the Russian revolution the Bolshevists were a faction of the socialist party of that country, subscribing to the doctrines of Marx and supporting a program of state socialism. They differed from other Marxian groups in their repudiation of the doctrine that socialism could be established only in countries with mature capitalistic institutions, a condition which Russia was obviously unable to fulfill, being a semi-feudal country. They believed successful revolution was possible in backward countries, if carried through by force at a time of crisis in political or economic affairs and maintained by a ruthless dictatorship using terroristic means. This group took the name Communist Party after their successful *coup d'état* in 1917. They immediately discarded their affiliation with the socialist movement and became bitter opponents of the type of socialism we have thus far been investigating, organizing opposition parties under their own name in all the principal countries and drawing this world-wide movement together under the leadership of the Third International, founded in Moscow in 1919. Today the conflict between communism and socialism is everywhere as bitter as the clash of either of these groups with their common conservative opponents.

This factional conflict within the revolutionary movement is confusing to the student who wishes to understand the essentials

of communism, for its program does not appear sufficiently different from that of socialism to account for so deadly a controversy. There were special reasons for the schism, arising from the attitude of orthodox socialists toward the World War and from incidents occurring during the Russian revolution; but these need not concern us. Reserving for later discussion the economic program of the communists, we may note that the distinguishing characteristic of this group is their method of revolution. They are scornful of the peaceful tactics of the orthodox socialists, believing that dispossession of the capitalist classes can never be brought about by these means and that to adopt them merely enervates the movement and commits it to a policy of flabby compromise. The communists base their strategy squarely on the class war, a succession of armed clashes with the forces of capitalism, culminating in a violent revolution on the Russian model, to be consolidated by a reign of terror. In harmony with this outlook they repudiate democracy, both in the organization of their party and in that of the revolutionary state, and propose a dictatorship of self-appointed leaders commanding a loyal well-disciplined proletarian army.[1]

Before outlining the program of communism, it would be well to warn the reader that this must not be assumed to coincide with the policies put into effect up to the present time by the Soviet Government in Russia. Internal and international conditions have caused that government to compromise with the original principles of communism. But quite apart from the influence of political expediency, the communist doctrine itself provides for an intermediate period between the seizure of the state and the final establishment of pure communism, during which the country will be carried through a stage of socialism not unlike that which the state socialists propose as their ultimate form of social organization. During this period, it is held, the ground will be prepared for a second transformation into communism.

The communist program. We must turn therefore to the theoretical writings of the communist leaders for a description of the ideal society proposed by this movement. In its extreme form the essen-

[1] These tactics have recently been modified for reasons of expediency in countries outside of Russia as described on a later page.

tial features of the communist program may be summarized as follows. There is to be collective ownership of all forms of industrial enterprise, administered by the government; labor alone is to give right to income, and all labor is to be performed for the government as the only employer; the entire social income is to accrue, in the first instance, to the government, to be distributed according to communist principles. The method of distribution differs from that of other socialist schemes. Wages are to be the same for all workers regardless of function and are to be paid in labor tickets which are redeemable in goods out of the general fund. Money is to disappear, and the only form of exchange will be that of "service for counter-service"; that is, labor service of the individual for the state and counter-service by the state in the form of consumption goods supplied to the individual. Many important types of goods — such as transportation, housing, light, fuel, perhaps food — are to be free for all citizens. Equality is to be the rule, and any departure from the rule is to be justified only on the ground of the greater need of the favored individuals. "From each according to his ability" is the formula governing the social service of the individual; and "to each according to his need," the formula determining his reward. The transition stage of the revolution is counted on to give the dictators opportunity, not only to bring the country's social institutions into conformity with this ultimate purpose, but to remould the behavior of the people and give them a motivation appropriate to a communistic society. When these ends have been achieved, dictatorship is to give way to a completely democratic system of government. Indeed the state itself, as a coercive organization, is to wither away and disappear as the people become motivated toward voluntary coöperation. The term "government," as used by the communists, connotes merely the executive organization of such a freely coöperating group.

The Communist campaign: Organization in Russia. The Communist Party of Russia provides the prototype of the movement's organization in all countries. Its essential features are: complete centralization of authority in the hands of the party's leaders; provision for fairly liberal debate of policy and tactics up to the time

when decision is taken as to the "party line"; ruthless suppression of all differences and the requirement of undeviating loyalty to the leaders when once decision has been reached. In other words, there is avowed dictatorship within the party as within the political state, which the party controls. Membership in the party is not a thing to be undertaken lightly. To join requires a confession of faith and devotion, a willingness to submit to arduous testing during a period of probation, and a readiness after becoming a full member to bear heavy financial and labor costs on behalf of the party's program. Members are taxed heavily in proportion to their earnings and are expected to give all their spare time to party works under the orders of their officers. Periodically the membership is "purged," through a comprehensive inquisition, which eliminates those of faltering faith and lukewarm works. Membership, in other words, is like that of a strict religious order, requiring zeal, devotion, and sacrifice. So it is in the parties of other countries. For this reason the Communist Party of Russia has never been numerically large, never including more than a trifling percentage of the people. But what is lost in numbers is more than made up by the discipline and drive of the movement. This quasi-religious aspect of communism, with the irrational and uncompromising attitudes which it evokes, is a salient feature of the movement in all countries.

The structure of the party is based on a nucleus or "cell," a local unit formed in a small geographical area, an industry, or other organization by three or more members. These local units are grouped into convenient regional organizations; these in turn form larger territorial groupings, which are brought together into the national party government at the top. Disregarding complexities of detail, it may be said that the structure is erected by electing delegates from the lower units to the higher, culminating in the election by the largest territorial sub-divisions of delegates to the All-Russian Congress, reputedly the supreme authority of the party.

But the actual lines of authority are concealed by this ostensibly democractic procedure. The Congress is an unwieldy body of large membership, meeting for short periods at long inter-

vals. It elects a Central Committee to meet more frequently on matters of policy, a Central Control Commission to serve as a supreme disciplinary authority over the party membership, and a Central Revision Committee to audit the party finances. The Central Control Commission, also a large body, delegates its authority to an elected executive board known as the Presidium. The Presidium, the Central Committee, and the Revision Commission join to elect the general secretary of the party and the Political Bureau, a small body in continuous session composed of the key men in the various central organizations. Through this complex hierarchy of interlocking organs supreme authority is gathered into the hands of the secretary and the Political Bureau. The secretary controls the paid party workers in all the local units and, using them as a personal following, becomes the leader of the entire party. It was by means of this position that Stalin first rose to dominance, and he subsequently achieved his present position of absolute personal dictatorship by using this machine to crush any of his former chiefs and colleagues among the older Bolsheviks who challenged his authority or differed with his policies. The Political Bureau is now composed of Stalin's closest associates. It is in reality the sovereign body in party government under the leadership of the Secretary, since it exercises all the powers of the Central Committee, exerts party discipline through the Control Commission, prepares and steers the program of the Congress, and controls the entire press and other publicity organs of the country.

International organization. So much for the national set-up in Russia. The same group of organizations which elects the secretary and the Political Bureau also appoints the delegates of the Russian Communist Party to the Third International, which attempts to impose centralized control over the policies and tactics of the party's branches in other countries. Communism is an international creed, ignoring national frontiers; the party is a single organization throughout the world, owing common allegiance to a single leadership in the Third International, which is dominated by the Russian branch only because of the latter's numerical superiority. This simple picture of a unified, world-wide communism is blurred by the fact that the movement in every other country than Russia

has been disrupted by bitter factional disputes. The International recognizes only one party in each country as orthodox and is itself composed of delegates from these legitimist parties, elected in proportion to their respective memberships. The dissident parties in other countries have usually split off from the main core of the movement as followers of one of Stalin's deposed rivals, most frequently Trotsky. They are bitter enemies of the present Russian leadership and of the legitimist parties in their own countries. These opposition factions work in harmony with one another and recently, under Trotsky's leadership, have shown signs of establishing a competing International. The orthodox factions however are everywhere in the majority.

There are some sixty national groups represented in the Third International, but it is impossible even to estimate the total party membership throughout the world, because of the secret character of the organization in many countries. The International is supposed by its constitution to meet every year; but following a few annual meetings immediately after its origin in 1919, the sessions have become less and less frequent. After the Fifth Congress in 1924, the next was postponed until 1928, and then no more were held until 1935. In this infrequency of meetings, as well as in the alterations of its announced policy, concerning which we shall speak in a moment, the International reflects the political exigencies of soviet Russia, particularly the recent desire of that country to avoid arousing the antagonism of other nations.

The pronouncements of the Third International, as the oracle of communism, are concerned far more with strategy than with doctrine. The basic doctrines of the movement are held in veneration as an inspired creed, revealed in its unalterable essentials by Marx and interpreted by later disciples, principally Lenin and Stalin. It is assumed, somewhat contrary to the fact, that these present-day interpretations are merely an exegesis, not an amendment, of Marx, and it is also taken for granted that the creed thus interpreted has the unquestioning adherence of all loyal members everywhere. The International has taken pains from time to time to expound to the world movement these interpretations of basic doctrine by the Soviet leaders, to justify by reference to the creed

the current policies of the Russian government, and thus to obtain the unqualified support of world communism for the régime in Russia. But since unanimity in the matter of doctrine is assumed, the task of the International has been principally that of issuing instructions on immediate tactics to the foreign branches of the party. In this it has shown itself to be dominated by the current needs of the Soviet Union. Its early meetings turned on the theme of class war and world-wide revolution; all branches of the party were encouraged to put themselves on a war footing, holding themselves aloof from all other movements however liberal, and fomenting conflict by guerilla tactics wherever opportunity appeared. As Russia became more and more absorbed in internal reconstruction, requiring stable commercial and political relations with other countries, and especially as the recent growth of Fascism brought its dangers of attack upon the Soviet Union and of suppression of communism in other countries, this line of strategy was altered. The orders issued by the Seventh Congress in 1935 disclose a complete change of front as contrasted with the early days. The communists of all nations are now ordered to drop their bellicose tactics and to coöperate with and if possible to join every liberal and progressive movement in their respective countries. Dimitrov, head of the International, defined this as "the Trojan horse policy." It is justified on the score that the communists must make common cause with the liberals against the growing power of the fascists, which is destructive to both. To be sure, communism as a method of government has less in harmony with liberalism than with Fascism, but the latter has issued a death sentence upon communism, while the liberal and progressive modern democracies, in accordance with their principles, are tolerant of it. Reference to these tactics as a Trojan horse policy refers to their ultimate objective. By "boring from within" the liberal organizations of society, it is hoped, not only to survive the present emergency created by Fascism, but eventually to capture these organizations and turn them to the long-run purposes of communism.

In America. This strategy, faithfully adopted by the American branch of the movement, makes it difficult to appraise the present importance of communism in the United States. The party was

founded here in 1919; four years later its 20,000 official members contained only 10 per cent who were English-speaking. In the 1924 and 1928 presidential elections, William Z. Foster, the party's candidate, polled 33,000 and 49,000 votes respectively. In 1936 the present leader, Earl Browder, polled 80,000 votes. In 1938, according to Browder's own statement, the total membership stood at 75,000, including those of the Youth League who were not full-fledged members. This is the paper record of the orthodox party, exclusive of the dissident factions which exist here as in other countries. As a legitimatist party this organization takes its orders from the Third International. Both in respect of structure and government and of the duties involved in membership, it is modelled upon the Russian party described above. Though of minute proportions numerically, its membership is active in many organizations which are non-communist — such as the Workers' Alliance, a union of the unemployed, numbering nearly 500,000, the American League for Peace and Democracy, claiming to represent some 4,000,000 members of radical labor and farmer organizations, some of the new CIO unions, and various radical organizations among the negroes, the tenant farmers, and the like. It is not to be inferred that communism has captured these organizations or is likely to do so. Nevertheless its influence through these affiliations is widespread, operating in the main through people who do not fully understand or subscribe to its doctrines. The official organ of the party in the United States is *The Daily Worker*.

Socialism in soviet Russia: Historical summary. The Russian Communist Party, described above, is not to be confused with the political structure of the Soviet Union, known officially as the Union of Socialist Soviet Republics. We shall not attempt a description of this complex governmental system nor discuss the political rights and obligations of the citizens under the soviet constitution. We are concerned here with the practical application to the life of the country of the economic and social principles avowed by the communist leaders. These principles are applied through the official organs of government. It suffices for our purpose to know that the governmental mechanism is in practice a dictatorship within itself and that it is completely within the control of the leaders of the

Communist Party. In dealing with the Russian experiment as an example of communist economics we must disregard the experiences of the years immediately following the revolution, for this period was too greatly disturbed by civil war and foreign invasion and by the profound economic collapse of the country to permit the communist leaders to concentrate on their program of social reconstruction. For the past decade however, and particularly since Stalin solidified his power, the dictators have held complete command of the situation, both within their own party and throughout the country at large, and the economic policies they have adopted during this time may fairly be taken as representative of communism in its early stages. A turning point of policy was reached when Stalin and his colleagues decided that their experiment could not wait upon successful class revolution in other countries but must proceed to establish socialism in Russia in isolation from the capitalist structure of other modern nations. To some extent the Russian program has been colored by the dictators' fear of attack from without. This fear in part accounts for the stress upon national self-sufficiency, the stimulation of patriotism in disharmony with the theoretical internationalism of the communist creed, and the large investment of capital and labor in military preparedness, which has retarded an elevation of the standard of living of the Russian people.

Planned economy. The nub of the soviet program inheres in a comprehensive economic plan, which embraces every phase of the national life and attempts to carry the entire social order toward pre-determined objectives. Certain phases of the plan are non-economic in character, dealing with such problems as the elimination of illiteracy, the promotion of education, hygiene, and public health, cultural activities, and the like. These are among the most interesting aspects of the plan, both in conception and achievement, but we must disregard them here for lack of space.

The planning mechanism is the State Planning Commission, created by the central government, together with its subordinate bodies for regional divisions and separate economic enterprises. The commission is composed of experts in the different departments of the national economy, dominated by party appointees who hold

the program true to the "party line." Its function is to prepare detailed schedules for every line of industrial and agrarian activity, covering the kinds and quantities of goods to be produced, together with the projected costs and prices, the rates of capital increase in various industries, the increase in the number and technical efficiency of laborers, and the like — the whole plan having relation to a term of years immediately following its adoption. During the operation of the plan, the Economic Council of the central government exercises supervision over it, checking current rates of production against the scheduled forecasts, and altering the schedules from time to time in accordance with changing political and economic conditions; the whole weight of governmental authority is behind the council's decisions. Planned economy in soviet Russia does not mean rigid, hard and fast adherence to the details of a program, but a flexible policy which, though outlined in advance and moving steadily toward pre-determined general objectives, allows for considerable readjustment along the way. Planned economy really dates from the adoption of the first "Five Year Plan" in 1928. A second "Five Year Plan" followed in 1933, and on January 1, 1938, the third "Five Year Plan" was launched. Soviet Russia has thus had a decade of experience with economic planning.

The principal features of these "plans" may be summarized as follows. First a determination to socialize the entire industry of the country under government ownership and operation, extinguishing private enterprise in all branches of industrial production and distribution. Second to establish socialism in an agrarian economy dominated at the time by millions of independent small-scale farms. These two great changes in organization were to result in the proletarization of the Russian people, creating a vast industrial army of wage earners, and giving the peasantry a similar status. Third to expand on a gigantic scale the industrial capital equipment of the country, to change its technique from small-scale to mass production and from handicraft to power-driven machinery, and greatly to increase its rate of output and lower its costs. This objective was concerned primarily with the heavy metal industries and with the sources of raw materials supplying them. Similarly to increase the output of foodstuffs and raw products of the sort required to make

Russia industrially self-sufficient. These were the major objectives. As will be seen, they involved a fundamental revolution in economic organization as well as an increase in productivity. The intention to raise the standard of living in the country by increasing the quantity and variety of consumable goods may be mentioned as a fourth objective. But the drive toward military preparedness, with its emphasis on the machine-making and war equipment industries, caused this phase of the plans to be set aside more or less completely. In the paragraphs which follow we shall attempt to summarize the result of the soviet program up to the present time.

Economic organization. Almost every department of the nation's economic life has been successfully transformed into a system of state socialism. The government owns and operates the railroads and other public utilities, the mines and forests, and all the factories; it monopolizes trade, both domestic and foreign; the banking and credit mechanism is in its hands. Agriculture is to some extent an exception. Under the soviet constitution legal title to all lands vests in the state, and in accordance with this principle there are vast farms owned and managed by the government. But the major form of agrarian organization is a transition stage between private enterprise and thoroughgoing socialism — a system of producers' coöperation, in which numerous individual peasants merge their land holdings into a single productive unit, which they manage collectively under governmental supervision. This system of collectivization has almost completely supplanted the former system of independent peasant holdings; in 1938 there were 243,000 collective farms, embracing 93 per cent of all the peasants and 99 per cent of the cultivated area. There is no private employment of wage labor even on the few remaining individual farms; the socialized industries provide the only source of wages. Small household industries are allowed to exist, provided they conform with the law prohibiting the hiring of labor and sell their products at prices fixed by the government. These vestiges of private industry are recognized as anomalous, their existence being justified by the government's inability thus far to provide an adequate supply of consumption goods through its own industries. The present rate

at which economic transformation is progressing indicates that reconstruction on socialist lines is virtually complete.

Growth of soviet industry. It is impossible to trace in detail the rapid development of industry under the soviet régime. At the outset Russia had only a rudimentary industrial equipment, particularly in the heavy metal and machine-making trades. Since 1928 she has vastly increased her capital equipment in these lines. Between 1927 and 1938 total industrial production increased in the ratio of 9 to 73, an eight-fold increase.[1] Many new industries have been established, notably in lines in which Russia was formerly dependent on foreign supplies; for example, tractors and other agricultural machinery, chemicals, and airplanes. From small beginnings Russia has advanced her production of iron and steel to a level now comparable with that of Germany. Electrification has proceeded at a rapid pace. The production of industrial raw materials has increased nearly four-fold during the ten-year period. In all these lines, basic to a highly industrialized society, the development of Russia under the régime of planned economy has been little short of phenomenal.

The record is not one of complete success. As stated above, the program for the consumption industries has lagged. Goods desperately needed for the everyday wants of the people remain very scarce and of poor quality. And beyond this there have been numerous shortcomings and failures in the industrial program, dramatized from time to time by the summary execution of the managers held responsible. Working with a labor supply wholly unacquainted with machine technique and with managers and technical advisers of little experience, the forced growth of industry has produced some awkward results. The new machine equipment has deteriorated rapidly through misuse and lack of repair; production costs measured in labor units have remained high; some of the output of the machine-making industries has been defective to the point of uselessness, and much of it has been of poor quality. The soviet authorities admit these shortcomings and attribute them to an over-ambitious program. The third "Five Year Plan," begun in

[1] The figures quoted in this section are taken from a special report of the Foreign Policy Association, *Foreign Policy Report*, June 1, 1938.

1938, calls for a slower rate of advance in most lines, and in some even for a decline of production below the figures actually attained the year before. This new policy of "deceleration" is made necessary, according to the soviet authorities, by the "incompetence, disorganization, and waste" which have developed under the severe pressure of the planned economy.

Labor and wages. The communist program is conceived in the interest of the wage earner, who is to be set free from the slavery which is his lot in capitalist society, given an independence which he has never formerly enjoyed, and accorded a larger share in the social income. In these as in other respects the experiment in soviet Russia has failed thus far to conform to communist principles. Labor policy has had an interesting development. At an early stage substantially equal wages were paid to all workers regardless of function, the higher types of expert and managerial labor being placed on about the same level of reward as the manual workers. This scheme broke down completely and had to be abandoned, due to its failure to stimulate the industry and efficiency of the abler workers. As the drive for rapid increase of productivity became more severe, the last vestiges of freedom and equality among wage earners were swept away. At the present time the wage system differs very little from that which obtains in capitalistic industry, except that the average real wage is very much lower and the extremes of inequality not so great. Piece-rate wages, payment in proportion to product, and higher rewards for services which are scarce are all characteristics of the soviet wage system. In like manner other socialist theories of labor management have had to be discarded. At first the powers of management were severely limited, and the workers were given virtual control over the operations of their respective industries. This system of industrial relations resulted in such hopeless confusion and inefficiency that the soviet authorities were obliged to reëstablish managerial autocracy in industry. In fact the control and discipline of labor in Russia is now much more complete and dictatorial than in even the most conservative of capitalistic nations. It is true that labor is organized comprehensively into trade unions, but these organizations are entirely subservient to the state and, as advocates of labor's welfare

on the job, are much less powerful than in capitalist countries. Strikes are outlawed as treason against the state. Even unintentional inefficiency in union management is severely punished, as in 1937, when the entire higher officialdom of the Central Trade Council was dismissed from office and four of the most prominent labor leaders were arrested. The unions no longer pretend to exercise any bargaining power in matters of wages, hours of labor, rates of output, or discipline. They are still allowed to function in adjusting minor local grievances, but on the whole they are used by the government to strengthen its drive for greater productivity and as agencies for administering the social insurance benefits.

The Russian people have always lived on a plane of poverty. It may be that the average real income is somewhat higher now than formerly, but it is still extremely low as judged by any standards known to modern industrial nations. In 1937 the government decreed a minimum wage equivalent to about 23 dollars a month for office and manual workers, a substantial increase over prevailing rates. During that year the average annual wage in industry was approximately 600 dollars. The program for 1938 sets the average annual wages in the major lines of industry at figures varying from about 517 to 741 dollars. These are the basic wage rates; under the bonus system the best workman can earn a good deal more. The prevailing inequality is shown by the legal rates themselves: an average of 300 rubles a month for the workers in industry, 10,000 rubles a month for engineers, technicians, and managers.[1]

These are money wages, the highest yet attained under the soviet régime. But they signify little in terms of real income. So great has been the shortage of the necessities of life that until recently these things were virtually unobtainable at the government shops, with the exception of food, which was distributed by a rationing system. The food ration was abolished in 1935, and since that time the people have had less difficulty in getting enough to eat. But the prices of most consumers' goods remain so high that the wage income for the average family is sufficient only to cover the bare necessities of life.

[1] The ruble has an official value equivalent to five cents.

In appraising the material welfare of the workers in the soviet régime, account must be taken of certain other benefits accorded them by the government. They are given security against unemployment, provided of course that they are willing to accept the terms offered by the only employer, the state. A comprehensive system of social legislation indemnifies them in case of accident or sickness. Maternity and child care, free education with stipends for students in the higher branches, and a certain amount of free medical service are also provided. For 1938 the government has budgeted a total expenditure of over a billion dollars for these free benefits to the workers.

Soviet finance. In the final analysis the costs of such planned economy take the form of sacrifices by the people of freedom and material comfort. This does not refer to the large-scale destruction of life, the exile or imprisonment of great numbers of people, the forced labor of political convicts, all incidents of a ruthless dictatorship. What is meant is that in its routine economic operation the program has fed on the toll taken from the standard of life of the common citizen. During this period Russia has been virtually unable to increase her resources by borrowing abroad. Yet she has imported vast quantities of capital equipment and raw materials and has drawn other large quantities from her internal resources. In the final analysis this has all come from the consumable incomes of the people, thus increasing the prevailing poverty. The government's power to determine the quantities of goods and services allowed to enter into direct consumption and the quantities to be employed in capital creation has forced the people to save, either against their will or without their knowledge.

But the day by day operation of the program has had to be financed, and the development of financing methods has produced some interesting results. In the first place, the soviet authorities soon had to abandon the communist doctrine that money is unnecessary in a socialist economy where the state monopolizes production and fixes wages and prices. The money system which has been set up in Russia is ostensibly on a gold basis, with a standard unit — the ruble — which represents a fixed weight of gold. But this standard monetary unit is entirely hypothetical and does not

exist in reality, being used merely as a unit of account for measuring values in government records and recording transactions in foreign trade. The money in circulation is irredeemable paper, whose value is determined solely by its purchasing power within the country. This value is not governed by the principles discussed in our earlier study of monetary theory, for the reason that prices of goods have been fixed by law and at times a rationing system has determined the distribution of goods among the consumers. In such a system there is no necessary connection between the quantity of money and the general price level. A given money income may fall into two parts: a portion which can be exchanged for a quantity of goods determined by authority, a remaining portion which has little value since there is scarcely anything to be bought with it.

In the second place, the financing of the planned economy has created property incomes, in contradiction of the socialist doctrine that all income must be drawn from labor The government, while obtaining what it could in purchasing power by taxation, has had also to resort to borrowing from its own people. Loans have been of two main varieties: deposits in the savings banks -- government institutions, whose cash resources have been invested in the industrial program — and bonds sold to the citizens High interest rates are paid, both on savings deposits and on government bonds, thus creating a new propertied class In 1936 the total internal loans of the Union amounted to 15 billion rubles; at the same time the accumulated savings of the people in the banks totalled some 7 billion rubles

Finally the development of the program has produced a situation in governmental financing which is illuminating as a prototype of a socialist economy. By 1937 the government had gathered into its own hands 95 per cent of the national income as compared with 27 per cent in 1928. This of course refers to "income paid out"; a considerable amount of real income still accrues directly to the people, particularly the peasants, in the form of products from small land holdings and the like. This large share of the national money income accrues, in the first instance, to the state by virtue of its monopoly of all enterprises producing for the market. Public finance under these conditions is chiefly an accounting of the

receipts and expenditures of productive enterprise. The government's budget records some income from taxes levied directly on the citizens, but in the main its receipts consist of portions of the income of the government monopolies reserved for purposes of the state. These are recorded in the budget as taxes, a "turnover tax" on sales, a "profit tax" on railways, factories, farms, etc.; but in essence all these exactions on the "profit" of business enterprise are subtractions from the incomes of the people. The "profit" is created arbitrarily by the power of the government to fix both the money costs and the sale prices of industry. What soviet public finance means in simple language is that the dictators fix the money wages which determine the people's income sufficiently below the prices of goods to leave a large surplus for their own use.

In 1937 the state budget totalled nearly 100 billion rubles. Governmental receipts from socialized industries accounted for 84 per cent of this total, taxes on individuals for 3 per cent, loans for 7 per cent, and miscellaneous incomes for the remainder. Among the expenditures the chief item was the financing of the economic plan — principally capital accumulation in new industries and in the improvement of existing industries — which accounted for 40 per cent of the total. Direct military expenditures absorbed 25 per cent, but the total cost of preparation for war was greater than this, since half of the current investment on the economic plan may properly be called indirect military expenditure on the war industries, transportation, etc. Educational, cultural, and social welfare activities of the state absorbed about 25 per cent of the expenditure.

LV

FASCISM. JUDGMENT OF RADICAL REFORM

Fascism and communism. It is a common opinion that Fascism as a social and political philosophy stands at the opposite pole from communism, representing conservative capitalist principles. The opinion has probably developed from the fact that in the countries — such as Italy and Germany — where Fascism has triumphed the movement arose as a belligerent opponent of radicalism, while in other countries, where Fascism is a minority movement, it is supported by ultra-conservative groups. In reality however Fascism as a political philosophy bears a close likeness to communism. They are both revolutionary movements opposed to democracy; both adopt the same technique of revolution — direct action by a firmly organized party under dictatorial leadership followed by terroristic methods. The Fascist conception of government by dictatorship and even its proposed organization of the political system appear to be identical with the ideas and practices of communism in these matters. Furthermore both subscribe to the principle of a planned and controlled economy, and both might well use their powers of governmental dictatorship in pursuit of the same objectives.

There are however many sharply defined contrasts between Fascism and communism, which become clear when one examines the two systems in actual operation. First there is a difference in the concept of the state, despite the fact that in both régimes the state is equally authoritarian. Fascism worships the state as a spiritual entity, superior to the sum of the people who compose it and eternal in its validity, while communism would eventually do away with the state entirely. As a corollary, Fascism is in theory militantly nationalistic, while communism is in theory international, recognizing no political frontiers. The fact that Russia has turned to nationalism is explained by the soviet leaders as a

temporary reaction to the present state of international affairs and is admitted to be contrary to their creed. There is no such disharmony between the rabid nationalism of Fascist countries and the basic doctrines of their leaders. Second there is a distinction between the underlying theories of government in the two movements. A Fascist state is to be dominated permanently by a small party, to which only the élite can belong, while communism, whose governing party is similarly organized in the early stages of the state, avows the intention of broadening its base until it includes everyone. Third the two movements differ in their conception of social classification. Communism not only believes in the class struggle as a historical fact but fosters it as a means of social evolution. Fascism, on the other hand, denies the historical necessity of the class struggle, explaining it as the artificial creation of demagoguery, and insisting that both proletariat and bourgeoisie receive their status from membership in the state and are bound by the same obligations of loyalty and service. Finally one finds a number of minor differences which it is difficult to reduce to formula. For example communism recognizes the essential equality of human beings, regardless of sex or race, granting to all the same legal rights and admitting them all to the same vocations, in the party, in the government, and in industry. Fascism, on the other hand, either excludes women from political and social activity or subordinates them to men, and it professes belief in extreme theories of racial superiority. There is a difference too between the attitudes of the two movements toward religion. Every loyal communist must be a professing atheist. Although the soviet government permits freedom of worship, it forbids church organization beyond the individual parish, thus weakening all organized religion, and it looks benevolently on the efforts of the party to stamp out religious belief by propaganda. Fascism requires no personal opposition to religion of its adherents, and it interferes with the churches only when they, as organizations, are hostile to its purposes.

Fascism in Italy: Fascist Party organization. The National Fascist Party was established in 1921 as the outcome of a congress held in Rome. Its original membership included individuals

drawn from all social classes and economic vocations, with middle-class people in the majority — about 150,000 in all. At the present time it enrolls some 8,500,000 members. The original structure was a makeshift arrangement of semi-independent local groups, organized for guerilla warfare. The structure and functions of the present party were defined by the Grand Council in decrees adopted in 1926 and amended in 1929.

By these official decrees the party is defined not as a political organization but as a civil militia serving the state. At its base is the local unit, the facio, under its own party secretary. The fasci are federated regionally, each such grouping being headed by the provincial secretary, and the provincial organizations are drawn together at the top in the National Directorate and the Grand Council, acting under the supreme command of Mussolini. The party officers are arranged in a hierarchy of subordination under the leadership of the Secretary-General, who is also secretary of the Grand Council. This post Mussolini does not hold; his title within the party is Il Duce, commander-in-chief.

Membership in the party carries with it professions of faith and obligations of service similar in nature to those exacted of its members by the Communist Party. Candidates for membership are given a searching examination, obliged to pass through a probationary period, and when admitted required to take an oath to obey "without discussion" the orders of Il Duce. Periodically the behavior of the members is examined by a disciplinary court, headed by the Secretary-General, and shortcomings in faith or works are punished to the point of expulsion from the party. Recently it has been decreed that recruiting of party membership shall be limited to the graduates of the Fascist youth organizations, which school the children in the creed of the movement and inure them to its discipline.

Relations of party to state. The Fascist Party and the official organs of government are so interwoven that in reality the party is the governing mechanism of the country. The Grand Council, governing body of the party, has been recognized by law as the coördinating agency of all departments of the state. Mussolini's dual position as president of the Grand Council and head of the

government — officially Prime Minister appointed by the King — has also been recognized by law.

The dominance of the party over the government can be made clear by tracing the process by which official policy is determined. Legislation originates with the head of the government; it goes for consideration to the Fascist Grand Council, is then examined by the ministry, and only after prior approval by these bodies is it presented to the Chamber of Deputies for adoption. The Chamber is the lower house of the legislature. The upper house — or Senate — which has little legislative function, is made up of members appointed for life by the King, a majority being Fascists. It is permissible too for the executive to issue decrees without consulting the Deputies, the decrees having the force of law for two years unless made permanent by subsequent approval of the Deputies.

The electoral process is under similarly rigid party control. The process originates in a list of one thousand candidates for the Chamber of Deputies presented to the Grand Council by the local party organs. From this list the council selects four hundred party members in good standing, whom it submits to the country for approval, the people voting on the list as a whole and not on the merits of the individual candidates.[1] Thus party and government are fused together through the person of Il Duce, legal head of both; through the constitutional powers of the Grand Council, supreme policy-forming body in both; and by means of party control of the membership and activities of the national legislature.

Fascist social organization. The "corporative state." Fascism has never developed a body of economic doctrine comparable in extent to that of communism. Its theories are stated principally in terms of the purposes or objectives which guide the dictatorship in its utilization of supreme power. The central objective has been to unify the social and economic life of the country so that all the nation's forces can be mobilized in the service of the state. There

[1] A decree of the Grand Council on October 8, 1938, indicates that the electoral character of the Chamber of Deputies will be brought to an end after April, 1939. After that date the Chamber will consist of all members of the Grand Council of Fascism and of the National Council of Corporations, who will hold membership *ex officiis*. The name will be changed to "Chamber of Fasces and Corporations." A description of the National Council of Corporations is given on a following page.

is no limit set to the extent of governmental interference with the economic activity of the citizens. Anything necessary to the political purpose of the moment may be done in the way of promoting, prohibiting, or regulating business enterprise, fixing prices, wages, and business costs, altering the details of industrial organization, or bringing the government directly into business activity as owner and manager of enterprises. In general however Fascism opposes outright governmental ownership of the means of production. It supports the principle of private property and relies on private initiative to function within the limits set by the state.

The structural organization of society prescribed by Fascist theory is known as the "corporative" system. Although developed with apparent precision in early Fascist writings, this system was not established legally until 1934, and it is even now incomplete. It has as its basis syndicates or unions of workers and employers respectively in each trade or calling, the two interests being obliged by law to organize separately. The state gives official recognition to only one such workers' and one employers' syndicate for each territorial unit — commune, province, or region — and the law provides that a syndicate, whose decisions will be binding on all workers and employers in the industry, can be formed by ten per cent of the individuals concerned. The local or communal syndicates of the two factions in each industry are grouped into larger regional organizations, depending on the geographical spread of the industry, and these in turn are formed into national federations. It is understood that the principle of separate but parallel organizations for workers and employers is observed throughout the entire structure. These parallel unions are then drawn together at the top into nine national confederations, eight of which comprise representatives of the unions of employers and workers in the chief subdivisions of industrial, agricultural, and commercial activity of the country, while the ninth represents the professional and artistic activities.

Here is an apparently symmetrical structure of capital and labor throughout the country. The principal function of the syndicates is to form collective agreements covering terms of employment, both in the local and regional divisions of the specific industry,

and nationally through the federations. Such agreements are binding on all employers and workers in the industry whether they have joined the recognized syndicate or not. Conciliation machinery operates to compose disagreements over the terms or the interpretation of a collective agreement; failing this, the case is submitted to a special Labor Court for final settlement, strikes and lockouts being prohibited. Since the confederations represent all factions in the industries, they are supposed to take a statesmanlike attitude toward industry as a whole, mapping out schedules of production, improvements in technique, and the like. In this they serve as a link between industry and government. They are dominated by political authority in that both the president and the council of each confederation are appointed by the government. In theory the syndicates themselves are voluntary organizations, but in reality all the syndical officials are designated by the Fascist Party.

On top of this industrial structure was erected in 1934 an all-inclusive system of "corporations," whose relation in function and authority to the syndical organizations is by no means clear. A corporate organization of Italian national life was envisaged in the original Fascist theory, and a National Council of Corporations had existed since 1930, but the constituent bodies of this Council were not created until four years later. In that year Mussolini announced the formation of twenty-two corporations, embracing every conceivable department of productive activity, each composed of representatives of employers, workers, and technicians in its respective field. Mussolini is president of each of these corporations, and members of the Fascist Party make up the rest of their officers The total membership of all the corporations, some 825 in all, form a Central Corporative Committee, which according to Mussolini's prediction will eventually displace the Chamber of Deputies and become the legislative organ of the state. Each corporation has a vertical structure, embracing all stages of production in its specific field from the supplying of raw materials to the completion of the finished product. At present the functions of the corporation are defined as follows: to act as advisory bodies to the government, to settle disputes between capital and labor,

and to regulate wages and other costs of production. But their future as forecast by Mussolini involves nothing less than absorption of all the powers of the state in domestic affairs.

This new-born and shifting system of social organization has not yet crystallized to a degree which would permit of precise description of the permanent structure of Fascist society. Thus far in matters of social organization, as in policy, the Fascist leaders have adopted opportunistic tactics, governed by the exigencies of the moment, setting up such mechanism in industry as will best serve their purposes. Still, assuming that the present corporate organization is to be considered permanent, and taking seriously what the Fascist leaders say as to their future purposes, the general nature of a Fascist society may be summarized somewhat as follows. The economic and political life of the people is to be organized in a single comprehensive set of institutions. Laissez faire is to disappear from both branches of the nation's life. Private initiative however is to be retained in economic affairs, production being left in the hands of organized groups of employers and workers, which will elevate private initiative from the plane of competing individual enterprises to the corporate plane embracing all branches of the industry. The state will intervene at will to regulate prices, wages, and profits, direct the flow of capital among different industries, and determine the relative quantities of different goods — all in the interests of maximum national power.

Avowed objectives of Fascism. As has been said, the Fascist dictatorship in Italy has never plotted in advance a course of economic program comparable with the planned economy of soviet Russia. Their policies have been largely opportunistic, adopted and altered in response to current situations, and guided as a whole toward a very few general objectives. These objectives have been in the main non-economic, although their attainment has involved serious economic consequences and economic means have been employed in furthering them. It is fair to measure the achievements of Fascism against its own avowed objectives, summarizing these for the purpose as follows: First a destruction of liberalism and democracy in the political life of the nation and the substitution of a totalitarian state controlled by the dictatorship of

a single party, itself governed by a dictatorial leadership. Second the promotion of national grandeur and prestige both in the domestic affairs of the country and in relation to other states. Third the accomplishment of these aims without essential change in the basic institutions of private property and individual enterprise in economic affairs.

Fascist achievement. The success of Fascism with regard to the first of these purposes has been made clear in the foregoing description of party and governmental organization and requires no further emphasis.

The second objective has taken concrete form, principally in terms of the military strength of Italy and the employment of military power and diplomacy to bring the nation to a position of importance in world affairs. But military preparedness has in turn necessitated comprehensive reshaping of the economic life of the country. At the outset Italy was largely dependent on foreign nations both for the instruments of warfare and for the raw materials — oil, metals, rubber, cotton — without which modern war cannot be carried on. With regard to raw materials this will always remain true unless a colonial empire containing these resources can be acquired, for the nature of the country has it so. In pursuit of their militaristic purposes the Fascists have come nearer than in any other phase of their program to the adoption of a rigid plan laid out in terms of years. They have undertaken to make the country as nearly self-sufficient as possible in war materials and within as short a space of time as may be. This has called for a huge investment of capital and labor in the war industries, such as shipbuilding, transportation, airplanes, the heavy metal trades, and other large outlays on raw materials, both in the accumulation of reserve stocks and in devising processes for the production of substitutes, such as new methods of refining oil from low-grade shale, the production of textiles from milk, synthetic products of the chemical industries, and the like.

Without offering any judgment upon the militaristic objectives of Fascism, the movement must be credited with remarkable achievements. Taking command of a country feeble in fighting strength, of slight influence in the international arena, and ap-

parently possessed of negligible surpluses of wealth and energy, the Fascists have brought Italy into the front rank among modern warlike nations. It is not necessary to support this statement with statistics of armament; its truth is demonstrated by recent diplomatic and military history. In achieving this result, Fascism has produced striking developments in Italian industry, not all of them pertaining directly to armaments. There has been for example a great expansion of the hydro-electric industry, evidenced by the fact that most of the large-scale factories now use electric power, that some 4,000 miles of railway have been electrified, and that electricity is used for lighting even in the small villages. The railways have been transformed in equipment and service until they now rank among the most efficient in the world. A large merchant marine has been developed; manufacturing enterprises previously existing have been expanded and modernized and new ones created. Although the major items in this industrial program have ultimate military purposes, they serve the peace-time needs of the country. In these respects, through singleness of purpose and centralized control, Fascism has shown remarkable power to mobilize the productive forces of the nation and to increase them by avoiding the waste and duplication of effort implied in a competitive system. It should be said perhaps that the economic efficiency of Fascist industry, according to the accepted test of production in relation to cost, has never been demonstrated, because of the established policy of promotion through government subsidies. Moreover the ultimate success of the whole military program has yet to be brought to the test of a war with a powerful state.

With regard to the avowed purpose of Fascism to preserve private property and private enterprise in economic affairs the record is inconclusive, since the whole system is still in transition. Outright governmental ownership and operation extends over only a small part of Italian industry, but rigid political control has robbed private property of most of its customary privileges and closed to private initiative most of its opportunities. Although the banking system is ostensibly a private enterprise, credit is a virtual government monopoly since both long- and short-term credit is available only to such borrowers, in such amounts, and at

such rates as the dictators direct. Industries are apparently organized in private ownership, but they produce what they are told in kind and quantity, and their costs are fixed by decree, as are many of their sale prices. When an enterprise established under private direction by political mandate proves unprofitable the government takes over the deficit. If a private business returns a profit, the government takes most of it in taxation. When wealth accumulates in private hands, expropriation occurs, as in the case of the recent capital levy of 5 per cent on landed and business property. Foreign trade and finance are completely monopolized by the government. Wage rates are fixed by political authority, and many items essential to the cost of living of the common people are sold at prices similarly fixed. In all but name this is a system of state capitalism.

Economic welfare of the people under Fascism. It is difficult to appraise the record of Fascism in terms of the material well-being of the people subject to it. Moreover such an appraisal is irrelevant to the avowed purposes of the program, for these do not aim at improvement in the economic lot of the citizen. The individual is supposed to gain enough in pride and self-esteem from the growing grandeur of the state to compensate for any sacrifices demanded of him. It must be remembered too, when measuring the economic returns of Fascism to the Italian people, that in addition to the heavy costs of her preparedness program Italy has to bear the cost of an expensive campaign in Ethiopia and is now drawing heavily on her own resources to modernize that country. There is no magic in dictatorship to prevent militarism and war from taking their inevitable toll from the real incomes of a people.

In some directions however Fascism has reacted favorably on the day-to-day lives of the Italian people. National pride, as well as military strength, has required that certain prevailing conditions of degradation and poverty be eliminated. Thus there has been a concerted attack on illiteracy, on unsanitary living conditions, on epidemic disease, and the like, all of which were symptoms of a low standard of living. In these matters the Fascist dictatorship has made notable advances throughout the country. Much has been done to bring under control the two great health scourges of Italy,

malaria and tuberculosis, to reduce infant and maternity mortality, to improve hygiene, and in matters of education to broaden the base of the school system, though of course under such strict political control as to prevent the development of intellectual freedom among the people. The progress of the country in health and hygiene is symbolized by a decline of the death rate from 13 per thousand in 1922 to 8 in 1937 and of the rate of infant mortality from upwards of 145 per thousand to 99. Significant also, as an indication of the elevation of family standards, is the decline of the birth rate from 30 per thousand in 1922 to 13 in 1937. The fact that this has occurred despite the effort of the government to increase the population through subsidies to large families and punitive taxes on unmarried people and childless unions is the more indicative of the response of the people to a new outlook on life.

In the more strictly economic aspects of their life the people have as yet gained little or nothing from Fascist rule. The level of material comfort among the Italian people was very low before the revolution. Nearly half the people lived by farming, either on small plots of ground owned by the family or by share-cropping the lands of others. This great section of the people never obtained more than a very simple subsistence from their labors, and both in relative numbers and in economic status the rural class of Italy remains today about where it was in 1922. The wage-earning class, on the other hand, has passed through various changes of fortune under the Fascist régime. For some of them — the higher wage groups prior to the revolution — there has been a sharp and permanent decline in real income. In fact during the first decade the average real income of the whole wage-earning class declined, but except for the aristocracy of the labor class, which has benefited less than others, some restoration of these losses has been made in the years just past. There is some evidence that the recent imperialistic ventures of the state have proved so burdensome to the propertied classes that Mussolini has been moved to seek a broader basis of support among the wage earners. Whether or no this is the correct interpretation, the government has recently adopted policies more favorable to labor. Since 1930 hours of work have been reduced by about 20 per cent; in many industries the forty-

hour week is now the standard. Money wages have been raised, and real wages, though still below their peak, are above their low point, despite the shortening of the work day in the meantime. One factor beneficial to all ranks of labor has been the reduction of unemployment. At the present time the average number of unemployed is about 500,000, less than 25 per cent of what prevailed in the lean years of the Fascist régime, and about equally superior to the average condition of the country prior to the revolution. This improvement in employment is due of course to the armament program and not to any distinctive merit of Fascist social theory, but its effect upon the welfare of the Italian people is none the less real. The price-fixing activities of the government, particularly with regard to foodstuffs, have recently tended to improve the real incomes of the common people. All in all, since the shift of policy referred to above, the dictators have taken care that their costly program should not bear too heavily upon the wage earners. The propertied classes, including the owners of industry, are in a less enviable situation; great private fortunes and large incomes have virtually disappeared under the régime of rigid governmental control of profit-making industry.

Public finance. The financing of government under the Fascist régime differs from that of soviet Russia in that there exists no such monopoly of productive enterprise from which to draw revenues. On the whole the Italian government relies, as do capitalistic governments, upon its power to draw upon the wealth and incomes of its subjects by loans and taxes. It has been a cause of wonderment to the rest of the world how so poor a country as Italy could support the extravagantly expensive policies of its dictatorial government, and question has been raised as to how long it can continue to do so. In the final analysis these policies have been financed at the expense of living standards within the country, for Italy, like Russia, has been unable to borrow abroad. Heavy taxes are laid wherever a surplus of private income can be found, directly on incomes and profits, indirectly on the whole population, but with greater relative severity on the poor through sales taxes. Wealth is subject to confiscation through capital levies. Government borrowing has been virtually in the form

of forced loans, directly from the banking structure, indirectly from the propertied classes and wage earners through the influence of propaganda and intimidation. Tested by any recognized principles of sound business, the financial condition of the government would be condemned as insecure. In addition to large floating indebtedness, the funded debt in 1937 totalled over 100 billion lira, and the government's budget contemplates enormous deficits for years to come. Under a dictatorship however these traditional signs of approaching bankruptcy do not necessarily bear that interpretation. The floating debt can be funded by decree, and the service charges on the funded debt can be reduced or postponed in an equally dictatorial manner, by altering the terms of the loan, by devaluing the monetary unit, or by a capital levy payable in the government's bonds — all of which devices have been employed in recent years. The limits of government solvency in such a régime are set only by the final refusal of the people to make further sacrifices in support of the dictator's purposes. There is no evidence that Italy has reached or is approaching that limit at the present time.

Fascism in Germany: Historical origin. We are concerned here only with those aspects of the Nazi movement which disclose its general principles of social organization and stamp it as a type of society contrasting in aims, methods, and results with democratic capitalism. The historical origins of the movement and the state of affairs in Germany which brought it to power in 1933 must be passed over for lack of space, as must also many of its most distinctive present policies, such as its racial theories, its cultural purposes, and its imperialistic program. The movement began with the formation of the National Socialist Workers Party in 1919 with a charter membership consisting of Adolf Hitler and six other men. This party entered the political arena, using the electoral machinery made available by a democratic constitution to gain adherents among the voters and representation in the national legislature. Its early tactics were not revolutionary, save in the sense that its members, organized on semi-military lines, indulged in sporadic violence to terrify opposition candidates and intimidate the voters. Its rise to power was through the ballot, the election of

1933 giving it control of the Reichstag with Hitler installed as Chancellor by constitutional means. Once in office however the Nazis used extra-legal means to consolidate their power, destroy all political opposition, and monopolize the agencies of government.

Political organization: The Nazi Party. The National Socialist Workers Party is both a political and a military mechanism. Its military arm, the brown-shirted Storm Troops (S.A.), was organized at an early date to give leadership to the rank and file of its membership and to supplement the propaganda of the leaders with acts of terrorism and violence. The Storm Troops are today a standing army of some 2,000,000 men, drilled and equipped; they constitute an armed force outside the official standing army of the Reich, which is an adjunct of the government, thus symbolizing that the party is itself a fighting as well as a governing machine. The élite of the Storm Troops form the Party Guard (S.S.), numbering some 250,000 armed and disciplined men, a sort of praetorian guard of the Leader.

As a political mechanism, the party is organized on the model of an all-inclusive national government. At its head stands the Leader and his cabinet of nineteen, each of whom is chief officer of a department or section of national policy. Some of these departmental heads have to do with party affairs, such as the Chief of Staff of the Party Police and the Leader of the Nazi Youth Organizations, who are members of the party cabinet. But included among them are departments which cover all the major subdivisions of the nation's government: Foreign Affairs, Colonies, the Treasury, the Judiciary, Commerce and Industry, Agriculture, the Press and Propaganda. Thus although the organs of official government remain in existence as administrative units, the party structure places over each of them a policy-forming body of its own. Each of the departments or sections has under its administration a complete functional organization extending throughout the country. The country is also subdivided regionally into sections, each under the control of a party chief.

In addition to its military and governmental structure the party includes certain affiliated organizations: youth leagues for boys and girls, leagues of professional men, welfare organizations, the influen-

tial Labor Front. Admission to membership in the party is now limited to graduates of the youth organizations, which serve to indoctrinate the children of the country in the National Socialist creed.

Like the state which it dominates, the Nazi Party is organized on the principle of leadership. Hitler as Supreme Leader appoints the members of the cabinet to serve not as his colleagues but as his subordinates. The structure of the party, the duties of its organs and officers, the policies it avows, are all dependent on his will. Each individual member of the party and its affiliated organizations swears unquestioning allegiance to him. A hierarchy of officials, each under complete subordination to an immediate superior, provides a channel through which this unlimited one-man authority extends from its center into every element of the national life.

The party and the state. In the course of the Nazi revolution the German Reich lost its federal character and became a completely centralized state. Local autonomies of every sort have been wiped out, and authority has been vested solely in a single national government. The powers of this government are unlimited, not only regionally but functionally, every constitutional guarantee of personal liberty or the immunities of property having been removed. There is no separation of legislative and executive power, for both are exercised by the same body of men. The legislative process itself may be set aside and the country governed by executive decree, each decree having the force of law for a four year period.

This authoritarian state is dominated by the party through devices which may be summarized as follows: The party is itself the only legal political organization of the people; all other political groupings in existence at the time of the revolution were disbanded, and the formation of new ones was prohibited. The citizens therefore can vote only for Nazi candidates; in fact an election is really only a plebiscite voicing approval of decisions already made. The state and the party are united at the top in Hitler, who is both Supreme Leader of the party and President-Chancellor of the Reich, holding tenure for life with power to appoint his subordinates. The bureaucracy of government from top to bottom is

appointive, with the Chancellor's approval required in each case. The party cabinet not only formulates policy before it is acted upon by the government, but some of its bureau chiefs are at the same time heads of the corresponding departments of government. A Nazi organization has been set up to scrutinize and control the actions of all governmental officials, so that in execution as in purpose laws of the state may be brought into harmony with National Socialist principles.

Nazi social and economic program. As its name implies, the Nazi movement began as an avowed enemy of the existing form of capitalism, professing socialist principles. Its official platform or program, first announced in 1920 and reaffirmed in 1932, contained many specific proposals to which either the socialists or the communists might well have subscribed, such as the abolition of all income not acquired from labor, the nationalization of all trusts and monopolies, the distribution of the profits of all large industries, the communalization of the retail marketing system, the expropriation without compensation of large landed estates, and a broad program of public care for the aged, the sick, the poor, the unemployed. In reality however these proposals were not informed with any definite social philosophy but by the existing miseries of the people. They were adopted to win adherents, as were other parts of the platform, many of whose policies were irrelevant or even contradictory to the social purposes stated above. It is a moot question whether the Nazis since coming into power have transformed the country into a socialist system or have left its basic capitalist institutions essentially unchanged. Some light may be thrown on this question by our subsequent review of actual policy under the Nazi régime. It seems a fair interpretation of the facts to state that in its formative years National Socialism was primarily a drive for political power, promising anything to an embittered and impoverished people which would be likely to win its support. Still its leaders did have in mind certain definite purposes to which political power would be put when once in their hands, and these may be called the broad social objectives of the movement.

The whole set of objectives coheres around a common point: the power and prestige of the German nation. Even the peculiar

racial and cultural doctrines of the movement, which we shall have to disregard, were in harmony with this major purpose. It involved the attainment of internal unity and stability, so that the whole energy of the people could be mobilized in its support. It involved the development of maximum military power for the nation, so that the Reich could regain its lost position of dominance in the family of nations. With these ends in view the internal program embraced the following policies. First the obliteration of democracy, with all its divisive influences in domestic politics, and the substitution of an authoritarian state under dictatorial leadership. Second the creation of a totalitarian social system; that is, an organization of society overriding all personal and property rights, which would bring the wealth and labor energy of the German people entirely into the service of the state. This meant of course unrestricted interference by the government in business enterprise, with power to destroy, regulate, or monopolize under government management every form of economic activity as occasion demanded. Third the promotion of loyalty and contentment in the great mass of the German people. This was to be accomplished in part by playing upon the emotions of racial and national pride which the humiliating position of Germany at the time had made extremely sensitive; in part by increasing the economic security, the dignity in the industrial world, and if possible the material comfort of the common people; and also, where necessary, by ruthless suppression and extermination of dissenting groups. It is this feature of their program which has required the Nazi leaders to seek the basis of their support primarily in the labor and lower middle classes. Finally the program included the utilization of these powers of political dictatorship and social control to focus the energies of the nation on military preparedness.

Nazi social organization: Government industry. Thus far in its development National Socialism has avoided any large measure of outright governmental ownership and operation of productive enterprise. The transportation industry, 80 per cent of which is public property, is an exception to the general rule. But this industry was already publicly owned long before the Nazi régime. It has proved necessary also to monopolize the sources of long-

and short-term credit in order to direct the flow of capital funds and the financing of trade; hence the Nazi régime has taken over the banks and the exchange markets. Through its control of exchange operations and by means of a quota system applied to imports and exports the state has come to dominate foreign commerce so completely that this too is sometimes called a governmental monopoly. But in this case the actual business transactions are carried on for the greater part by private concerns subject to detailed restrictions, rather than by governmental agencies, so that technically it is not a state enterprise.

The "estates." The great bulk of Germany's production enterprise remains therefore in private hands, but in harmony with their totalitarian principles the Nazi leaders have regimented each department of economic activity into a nation-wide organization, called an "estate" These in their broadest groupings are two in number: the Estate of Industry and Commerce and the Estate of Agriculture. The organization of each estate is mapped out on paper down to the smallest detail. The Estate of Industry and Commerce, for example, is subdivided into six separate functional entities: manufacture, commerce, banking, insurance, handicraft, electric power. It is also subdivided regionally into eighteen district boards. Each functional subdivision is itself broken down into smaller units, representing separate categories of productive enterprise. The functional and regional subdivisions have their own governing bodies, and at the head of the estate stands the Reich Economic Chamber, a policy-forming body for all the different branches of industry and trade. Agriculture is similarly organized into functional and regional subdivisions, with a complexity of governing boards and councils drawn together at the top into a national chamber. Membership in one of these estates or in one of its appropriate subdivisions is compulsory for all business enterprises, large or small, throughout the Reich.

It would not be profitable to describe in detail the elaborate organization of Germany's business activity under the Nazi régime. The whole system is in a state of flux, with new governing agencies appearing and disappearing, and areas of jurisdiction shifting from one agency to another as Nazi policy develops. It

appears true however that the general principles and purposes which this form of compulsory, non-competitive, nation-wide organization embodies are firmly established in German economic life. The details of organization may change, but the farmer, the merchant, the manufacturer, the banker — in fact any person seeking an income from the ownership or management of an enterprise — has accepted the principle of subordination in a structure dominated by the state. These organizations are supposed to provide for self-government in industry. In reality they are all founded on the leadership principle, the government having power to appoint their officers and to prepare policies determining the kinds and quantities of goods to be produced, the methods of production or business operation to be employed, the costs to be met, and the sale prices to be accepted. In preparing these policies the government may draw in from the elaborate structure of the estates representatives of the business world to serve in a consultative and advisory capacity. But the final decisions are political and authoritarian, and the machinery of the estates is used to enforce them upon the individual businessman.

Labor organization. Labor is not included in the organizations described above, these being limited to the property and managerial interests. All previously existing trade unions have been dissolved, and in their place has been set up a nation-wide grouping of both employers and workers — the so-called Labor Front. Membership is not legally compulsory but, in view of the dangers encountered by the dissenter, is virtually so in fact. The basic unit includes the employer and the wage earners in a single concern; these are gathered into fairly small local groups, some 15,000 in number, which are drawn together into some 800 district groups, and these in turn into about 30 regional divisions. The whole organization is political in purpose. Its units from top to bottom are affiliated with the National Socialist Party, its officers must all be party members, and it is governed from the top by the cabinet minister of the party in charge of the Department of Political Organization. The government is of course dictatorial rather than democratic.

The Labor Front has no bargaining functions but exists primarily to unite employers and workers in broadly cultural activities and

in numerous coöperative undertakings for the betterment of living conditions among the common people. It administers industrial insurance, savings, and building companies. Its "Strength Through Joy" society provides travel, sports, and other recreation for its members. Its educational branch supplies reading material to the working classes and organizes literary, debating, and artistic societies. These activities, though not directly affecting the employment contracts of the wage earners, have had a profound emotional effect upon the common people of Germany, giving them a feeling of self-respect and importance in the scheme of things which has won their loyalty to the Nazi régime.

The outlook of the Nazis repudiates the Marxian doctrines of social stratification and class struggle. In accordance with this view strikes and lockouts are prohibited as offenses against the state, and workers and employers are obliged as a matter of patriotic honor to come to an agreement as to the terms upon which they are to collaborate in service to the nation. Courts of Social Honor have been established to pass on charges brought by one side or the other involving abuses of power or recalcitrance in promoting the maximum efficiency of the enterprise. These courts however can only punish offenses; they do not decide the terms of the employment contract. This important function is performed in the first instance by a peculiarly Nazi organization of the labor market on leadership lines. The employer of each concern is designated the "leader" of that enterprise, with powers to announce the terms of employment, having first consulted a council made up of representatives of the wage earners, called his "followers." The council members, though elective, must be selected from a panel of candidates drawn up by the leader with the approval of the local unit of the Nazi Party, one such unit having been organized in every business concern of the country employing more than twenty persons. This whole procedure is in turn supervised by a system of regional Labor Trustees, whose personnel is appointed by the government, and who report to and receive instructions from the Ministry of Labor. The trustees have virtually unlimited power. They do not interfere in a concern's labor policy as long as this conforms to the requirements of the Nazi program;

but they may at any time abrogate a labor contract and fix new schedules of wages and hours. An important item of their authority is the power to forbid any reduction of the work force by more than ten per cent, regardless of the circumstances. In reality therefore the labor costs of industry and correlatively the money incomes and work conditions of the wage earners are determined by political authority.

A form of state capitalism. Fascism, of which the Nazi program is an expression, is commonly believed to be a conservative movement, defending capitalism against radical movements, particularly communism. National Socialism in its rise to power in Germany did have the moral and financial support of the propertied interests, especially the large land owners and the industrialists, who viewed it as an ally in their fight against a rapidly growing communist movement. Nevertheless our brief review of social organization in Nazi Germany has shown that in that country, as in Italy, Fascism is in fundamental conflict with the essential institutions of capitalism. Private enterprise remains in existence and private initiative is still relied upon to keep the productive system in operation, but property has no rights immune from political authority, and private initiative remains undisturbed only so long as it obeys the orders of government. On the other hand, although the base of the movement is in the wage-earning and lower middle classes, they too have no rights as economic factors and no self-determination as persons. If it is important to attach a descriptive name to the Nazi social system, it is most accurately described as state capitalism.

Economic results of the Nazi program. The period of the Nazi rule has been too brief to permit of any final appraisal in economic terms. During this time every phase of policy has been dominated by militaristic purposes, which have entailed immense investments of capital and labor in lines not productive of consumable income and the equally large sacrifices in the interests of national self-sufficiency which militarism postulates. In this major purpose the Nazi régime has been successful at least in so far as the immediate fighting strength of the nation is concerned. The rapid expansion of every branch of Germany's war machine in size, in material

equipment, and in organization attests the efficacy of the Nazi principles of social unity and leadership to develop productive power and keep it focused. The drive toward economic self-sufficiency has not attained its goal and will probably never do so, because of insuperable natural difficulties. It is in this branch of their program only that the Nazi régime has set up an economic plan on the Russian model — a "Four Year Plan" for the attainment of self-sufficiency in food and raw materials launched by Hitler in 1936. Thus far no important results have been obtained from this plan. Isolated instances of the discovery of synthetic substitutes for natural products have been dramatized in the press, but the food situation in Germany is if anything worse than it was before, and the country remains as dependent as formerly on foreign supply of raw materials.

Having in mind these costly aspects of the Nazi program, it is understandable that the people have not benefited in terms of economic welfare. Real incomes have declined. There is a shortage of goods required by the common people, which has forced the government to ration supplies. This is particularly evident in the case of foodstuffs, which have fallen so far short that the standard of nutrition for the whole country has been estimated to have declined by nearly twenty per cent. The Nazi régime has adjusted its policy to this situation in two ways: by preventing the dearth from enriching the merchants and middlemen at the expense of the consumer, and by bringing privation to bear with greater relative severity upon the wealthy classes than upon the common mass of the people. The marketing system has been subjected to such meticulous regulation that it remains a private enterprise only in form. Wholesale trade in food has been abolished entirely, and the profit margin of the retailer has been drastically reduced by setting minimum prices on his purchases from producers — designed to promote production as a phase of the self-sufficiency program — and maximum prices on his sales to consumers. The latest development of the marketing policy, not yet universally applied throughout the country, is the system of "monopoly shops." This assigns to each retailer of food a list of customers who can buy only from him and to whom alone he can sell. This

system has the effect of assuring the merchant a definite turn-over and a calculable income, which he can neither increase nor decrease. It has the more important result of bringing the rich down to the level of the poor in respect of articles of consumption whose supplies are scarce, since all buyers, regardless of their incomes, can obtain only an official ration from a single designated shop. The price-fixing policy also furthers this effort at equalizing living conditions among the various social classes by discriminating in favor of ordinary as compared with "quality" goods. In sum the declining prosperity of the country has produced a greater equality of real incomes, while reducing the average to a lower level.

In estimating the effect of Nazi rule on the common people, one benefit must not be overlooked -- the elimination of unemployment. When Hitler took over the government, unemployment was a major national problem, not only by reason of the immense number of people actually out of work at the time, but because the workers everywhere were affected by the prevailing insecurity. The Nazis first solved this problem by direct action of government. The policies first adopted were of a type which called for a maximum of labor and a minimum of capital equipment, such as road building, land reclamation, reforestation. The unemployed were enrolled in labor armies and put to work in these undertakings under quasi-military command. As the rearmament program gathered momentum, expanding the operations of the German industries, private employment was increased sufficiently to reduce the need for labor armies and make-work projects. Today every German wage earner, if loyal to the régime, is virtually assured of a job. Even temporary unemployment is reduced by the power of the Labor Trustees to prohibit any large reduction in the work staff of a private employer until jobs can be found elsewhere.

Immediately beneficial as this situation is to the working classes, it is fundamentally precarious and may in the long run react disastrously upon them. The present volume of employment is maintained in the final analysis by the armament program. Stripped of its complexities, this program involves taking present incomes by taxes, loans, and other exactions and transferring them

to labor in the war industries; that is to say, a part of the German people is required to support the rest. The essential feature of the situation is that the workers thus drawing support are producing virtually nothing which will increase the future income of the country. The arrangement, in other words, does not create the means from which workers in the future can draw their incomes and so does not provide for its own elimination from the nation's economy. How long it can continue depends on the success of the government year after year in finding and appropriating surpluses of income from its subjects. The Reich has published no budget since 1934, and it is therefore impossible to discover how great a load of debt and current deficit the program has piled up. But that the process cannot go on indefinitely is certain, and when it comes to an end it will be found to have provided no permanent solution to the problem of unemployment. In the meantime the whole business structure in Germany, even in lines not directly identified with the war industries, has been distorted into harmony with the militaristic program, thus making the problem of subsequent readjustment more grave and the danger of eventual collapse more general.

General appraisal of radical programs: Initial qualifications. In this and the immediately preceding chapter we have considered the abstract principles underlying the socialist program and such general principles as can be discovered implicit in Fascism. We have also sketched briefly the concrete policies and the accomplishments of the only thorough-going experiment in socialism with which the world is familiar, and we have done the same with regard to the two Fascist régimes. It would be in point now to estimate the merits of these radical reform programs as compared with capitalism, so as to reach a final conclusion as to their relative desirability as forms of social organization. However to attempt this in any comprehensive way would lead us into generalities to which so many exceptions could be found in the actual policies of the three countries that our conclusions would have little final validity. Communism and Fascism cannot without qualification be lumped together in contrast with capitalism, for they have different aims with regard to capitalist economic institutions. On the other hand, they cannot be presented as antitheses of each other, since they are

identical in one important respect; namely, their method of revolution and social control after revolution. Furthermore communism and socialism cannot be lumped together, because the methods and many of the policies of communism are anathema to the majority of the socialists. This complexity of similarities and contrasts must be borne in mind in the discussion which follows.

Authority versus liberty. As has been said, communism and Fascism have one essential feature in common; *i.e.*, their destruction of individual freedom, beginning in the political realm and extending into every department of human activity. This is not an accidental effect of their program but is a matter of avowed conviction. Now one of the chief charges in the indictment of capitalism is that this system of society assures no real freedom for the average man. Granting this accusation all the weight it deserves, it certainly remains true that the evil has been magnified in the systems which have been substituted for the capitalist order. We need not dwell upon the incidents of massacre, imprisonment, and brutality which have horrified the democratic world, these are characteristic of dictatorship regardless of its social philosophy. But it is of the essence of capitalism, with all its inequalities in the distribution of wealth, power, and privilege, to rely upon the voluntary actions of individuals to provide for social needs. And it is equally essential to any controlled economy to substitute compulsion for voluntary action. The modern world now has experience of both methods in operation simultaneously and can easily judge of their results in terms of human freedom.

The constitutional socialists will hold this criticism inapplicable to their system, since they support democracy in industry and in political government. But it is questionable whether in practice they could avoid a wide application of compulsion. Even in a democratic régime a system of public control, instituted over the major economic activities of the citizens, must ramify until it embraces practically all the interests of a normal human life. To most men a society where the type of work in which a man may engage, the income which he receives, and the kinds of wants he is permitted to satisfy are all determined by authority bears little resemblance to freedom.

These radical systems, repudiating as they do self-interest as the motive force in human action, have been obliged to play upon some other motive to bring the wills of the people to support the commands of the state. In the final analysis mere compulsion, exerted through fear and intimidation, has not proved effective in calling forth the productive efforts of the people. In its theoretical literature socialism paints a picture, very appealing to all men of right mind, of a society from which acquisitiveness and the pursuit of gain will have vanished and men will be driven by loftier motives of brotherhood to work industriously for the common good. Socialism in practice has indeed found a compelling motive to substitute for individual self-interest, and so has Fascism. But this motive in soviet Russia is not the noble impulse of human brotherhood which socialism envisages, it is in fact the same motive which impels the peoples of the Fascist countries to respond to the commands of their dictators, namely, militant nationalism. It is true that as developed by communism, patriotic fervor has not displayed all the excesses of Fascism. Nevertheless it is the same sort of nationalistic pride, colored with suspicion and contempt of other peoples, upon which Stalin plays in arousing the popular will. There is nothing in the record of belligerent patriotism, with its past history and present threat of war, to commend it as a superior method of organizing the world of humanity. Nor is its effect upon the character of the individual human being demonstrably superior to that of self-interest, corrosive as is the latter in its grosser forms.

Capital formation. There are certain inferences of a general character to be drawn from the actual processes of a planned and controlled economy. One of these has to do with the process of capital formation in a socialist régime. The incentive of an interest return is believed to be necessary for the accumulation of capital in a system of free enterprise. The authoritarian experiments have shown that capital can be provided under authoritarian management without appealing in this way to individual self-interest. The soviet government in particular has certainly acquired a vast capital equipment. There has been some resort to borrowing through interest-bearing government bonds; but the funds raised

in this way have been only a small part of the total capital accumulation, and the purchase of the bonds by the people has been more largely a response to social pressure than to the attraction of the promised interest payment. For the most part the capital has been provided by the people through savings which they either did not know they were making or which they could not avoid. Funds raised by taxation and goods acquired by governmental requisition are examples of compulsory savings. A more general form of saving, often unknown to the people who have borne the burden, has resulted from the government's power to determine the quantities of goods and services which were allowed to enter into direct consumption and the quantities to be employed in the production of capital instruments, a power which it possessed by reason of its control of the entire productive system. To the average man in Russia the effect of this governmental policy has appeared in the form of a scarcity of consumption goods; he need not know that this scarcity merely reflected the government's decision to accumulate capital as rapidly as possible or, in other words, that he was bearing the burden of saving. It is in these ways that a socialist régime, if given sufficient authority over the people, can meet its needs for capital equipment while dispensing with the incentive of personal award to the saver.

But in connection with this matter it should be noted that a socialist system has no power to alter the process of capital formation in its essentials. The accumulation of capital involves a sacrifice of consumable income by some individual or group in society. Socialism does not make saving burdenless but merely conceals the burden. This concealment is possible because the individuals who make the sacrifice do not take part in the decisions of the governing authorities. In other words, it is the dictatorial character of the soviet government which gives it power to accumulate capital without the knowledge or against the will of those who provide income for this purpose. In a democratic system the people at large would be called upon to make the decision whether to live more comfortably in the present or provide for the increased comfort of later years and succeeding generations; and it is a safe conjecture that they would require some real incentive to induce them

to forego present enjoyment. The soviet leaders foretell the time when their governmental régime will be thoroughly democratic. Until the matter is put to a test at that time we have no evidence that democratic socialism — and this is the avowed objective of all socialist systems — will be able to acquire and maintain its indispensable capital equipment while abolishing pecuniary rewards to savers.

The planned economy. A similar observation may be made regarding the effectiveness of a planned economy as compared with that of free enterprise. Since comprehensive planning has been put to the test only in Russia, inferences regarding it must relate primarily to that country. It would appear that the soviet experiment in economic planning has been, on the whole, a success. It is true that things have not gone always according to prediction. There have been partial failures and unforseen difficulties in many branches of the program, but this is to be expected of so gigantic and complex an undertaking. The planning commission has succeeded in guiding the nation's economic activity in general toward the goals set up in advance, and this achievement has been interpreted as evidence of a superior virtue in communism or socialism. No such inference can be drawn from the facts. The success of the experiment in economic planning has been due to the dictatorship whose power stands behind it; and any equally strong dictatorship, whether of a socialist or a capitalist variety, would be able to accomplish similar results. Recalling again that communism in its true form presupposes a régime of democracy, the Russian experiment gives us no assurance as to the practicability of this aspect of the communist program.

Furthermore there is reason to believe that the attempt at economic planning has been greatly facilitated by certain special conditions in the country. This phase of the matter can best be made clear by stating that throughout this period the country has been laboring under conditions of extreme scarcity in all departments of its economic life. There has been a dearth of virtually every type of good and service, both for the satisfaction of human needs and for the development of basic productive equipment. In many respects such a situation greatly simplifies the task of a planning

commission. There is no problem of oversupply in any line of production; the authorities can safely assume a consumers' requirement larger than any product that it is possible for them to provide. There is no danger of an excess of labor in any branch of activity, since all are alike under-supplied with labor. No great thought need be given to the question whether the people would prefer more of this good and less of that, or would desire different styles or qualities of goods, since in the prevailing scarcity they will be happy to have any increase at all in the provision for their needs. Obviously this condition is not typical of advanced industrial societies, where the problem is one of relative overproduction in many lines of activity and relative oversupply of certain types of labor and capital, and where the preferential demands of the people must be considered. In these more typical situations the task of a planning commission must be immensely more complex and difficult, and until the planned economy has been tested under these more typical conditions it is unsafe to pass favorable judgment upon it.

What has been said applies in some measure also to economic planning in a Fascist régime. It is unthinkable that anyone seriously desires the nations of the world to remain perpetually on a war footing, their entire life dominated by the voracious demands of militarism. What success economic planning has had in Germany and Italy is pertinent almost exclusively to war. But the superiority of a planned economy as compared with one of individual enterprise can be decided finally only in terms of the welfare of the people in peace time. Neither communist nor Fascist experiments give us evidence on this point.

Judgment of socialism. In conclusion it will be profitable to return to the socialist indictment of capitalism as summarized in the preceding chapter. Every thinking man admits that each of the charges in this indictment contains elements of truth. The present order is no more perfect than is any other work of man. In fact it contains many and serious faults, faults of inefficiency, of injustice, above all of retarded and imperfect adjustment to the changing conditions of human life. The socialist movement is to be credited with the great service of having focused attention on these faults and accelerated their correction. Its own program however is not

convincing. In addition to specific faults discovered in our examination of socialism in practice, the case against it may be summarized under three principal heads.

The indictment of capitalism. Socialism exaggerates the inefficiency of the present order. The indictment of capitalism is overdrawn. After all has been said that can be said against it, our system has succeeded in maintaining the largest population the world has ever known upon the highest level of average comfort. It has crowded into the short space of a century and a half greater advance in all the arts of production than was achieved during the whole preceding space of society's history. It has equipped mankind with an unprecedented supply of capital instruments and is increasing the quantity and improving the quality of this equipment without pause. The uncontrolled individualistic system of production and marketing, despite grievous lapses in times of economic unsettlement, has worked for many years to supply the individual consumer with want-satisfying goods drawn from the resources of the entire world. It has accomplished this by providing incentives which call into play two powerful human impulses: acquisitiveness and the desire to excel in competition. The institution of private property supplies the former of these incentives; individual enterprise, the latter. Socialism aims to suppress the motive of acquisition entirely and to deprive success in competition of all material rewards. In their place it would rely upon the motives of brotherhood, social service, and altruism. It is not necessary to deny that human nature possesses these latter characteristics nor that they are of higher order than those nourished by a capitalist environment. But there is grave reason to doubt that they will prove as effective in stimulating human energy along lines of economic activity as the motives which have been responsible for the unparalleled material success of capitalism.

The vitality of capitalism. Socialism underestimates the vitality of capitalism. The survival of social institutions depends upon their power of adjustment to new conditions as they arise and of providing within themselves mechanism for solving the problems which are created by their own growth. Capitalism has this power. It is a flexible system. The institutions of property and competition

can be qualified when need arises without destroying their fundamental utility. Collective bargaining can be substituted for the individual contract when the latter works injury to the weaker members of society. Social insurance can be developed to safeguard the workman against many of the hazards of the wage earner's life. Labor legislation can assure him minimum standards of safety, sanitation, even of income. Collective provision may be made for the services which private initiative will not supply, such as education, agencies of public health, recreation facilities. Producers can be restrained from undermining and destroying each other by unfair competitive practices. The consumer can be protected against adulterated goods and scrimped quantities, by laws enforcing standards of purity and honesty of weight and measure. The overweening power of great business can be restrained by public control. All these things can be done without discarding the basic institutions of society. Although we may not be content with the rate of progress attained along these lines in the past, that there has been progress no one can deny. The true conservatives of our day are those who lend their support to these reforms, those who oppose timely readjustment of traditional ways and methods must be ranked among the enemies of the present order.

Practical difficulties of socialism. The socialist program creates grave practical difficulties. All forms of socialism impose on some public body the task of determining the prices of different goods and services. Is not this a task too great for human intelligence in so complex a society as ours? Price fixing involves a determination of quantities to be produced and a rationing of individual consumption. After wages have been fixed for all classes of labor, the state or some other public body must then see to it that society is supplied with the right amounts of each type of worker. When a price has been set upon wheat or coal or shoes, some authority must then supervise production and consumption. These difficulties multiply as one ponders the problem of authoritative control. The Russian experiment, for reasons already given, provides no proof that a politically controlled economy is practicable under normal conditions.

Again some method must be devised for detecting the natural leaders of men, advancing them to the positions for which they are best qualified. and giving them the requisite authority to make their powers effective. Each of these steps creates serious difficulties. For the selection of leaders democratic socialism has to propose no better method than that of public competitive examination. This always works against the innovator in favor of the man who conforms to the traditionalism so dear to the hearts of public officials. It places a handicap in favor of the politician, the time-server, the sycophant. The adjustment of capacity to task is no less difficult. Today it is worked out by process of trial and error. In the society of the socialists it is apt to be decided by rule of thumb. In this connection it must be remembered that socialism postulates a democratic order of society. Democracy has always been suspicious of concentrated authority. Would discipline be improved if each manager owed his tenure of office to the good-will of his subordinates? It is not enough for socialists to argue that these problems are not now perfectly solved. They must propose a superior method.

Thirdly socialism must provide for the maintenance and improvement of the capital equipment. No possibility would remain for the saving of capital by individual initiative, since interest would be destroyed as a private income. The Russian experiment has shown that it is possible for society to provide a different method of saving. The state, as recipient of the social income, can withhold a portion to be reinvested, thus forcing the people to save against their will or without their knowledge. But in view of the impatience of men for immediate income and of the subservience of office holders to popular clamor, is it likely that any group of public officials could be found with sufficient courage and foresight to administer such a system wisely? As regards the improvement of technical processes, it is argued that these come in final analysis from scientists who labor without thought of material reward for the sake of their love of truth. This may be granted without removing the difficulty. Today the discoveries of science are adapted to the economic service of mankind, not by the discoverers, but by other men who are certainly intent on personal reward.

Pure science might flourish under socialism, but invention would lag behind because of the apathy, the obtuseness, or the abhorrence of change of public officials.

SUGGESTIONS FOR FURTHER READING ON PART IX, PROPOSED CHANGES OF THE ECONOMIC SYSTEM

I. Profit Sharing and Coöperation

Balderston, C. C. *Profit Sharing for Wage Earners.* New York, 1937

Barritt, A. W., *et al. Profit Sharing.* New York, 1918

Bowie, J. A. *Sharing Profits with Employees.* London, 1922

Heilman, R. E. *Profit Sharing.* New York, 1918

Brainerd, J. G. *Consumers Cooperation.* Philadelphia, 1937

Gide, Charles (edited by Cedric Long). *Consumers' Cooperative Societies.* New York, 1922

Warbasse, J. P. *Coöperative Democracy.* New York, 1927

II. Socialism and Communism

A. *Theoretical*

Marx, K. (edited by Algernon Lee). *Essentials of Marx.* New York, 1926

Lenin, N. *Imperialism. The State and Revolution.* New York, 1926

Laski, H. J. *Communism.* New York, 1927

Laidler, H. W. *A History of Socialist Thought.* New York, 1927

Russell, B. *Prospects of Industrial Civilization.* New York, 1923

Skelton, O. B. *Socialism, a Critical Analysis.* Boston, 1911

Tawney, R. H. *The Acquisitive Society.* New York, 1920

Pigou, A. C. *Socialism versus Capitalism.* London, 1937

B. *Descriptive*

Batsell, W. R. *Soviet Rule of Russia.* New York, 1929

Cole, G. D. H. *Guild Socialism.* London, 1920

Hoover, C. B. *Economic Life of Soviet Russia.* New York, 1931

Russell, B. *Proposed Roads to Freedom.* New York, 1919

Webb, S. and B. *A Constitution for the Socialist Commonwealth of Great Britain.* London, 1920

—— *Soviet Communism.* London, 1937

McCrary, S. *The National Socialist State.* Dallas, 1935

INDEX

DELHI POLYTECHNIC
LIBRARY
CLASS NO. 311
BOOK NO. W 47 B
ACCESSION NO. 8716